## SEQUENCE LISTING <110> Lyamichev, Victor <120> CHARGE TAGS AND SEPARATION OF NUCLEIC ACID MOLECULES <130> FORS 4912 <160> 85 5 <170> PatentIn version 3.0 <210> 1 10 <211> 23 <212> DNA <213> synthetic 15 <220> <221> misc\_feature (1)...(2)<222> 20 <223> misc. feature <400> 1 23 tntcttttca ccagcgagac ggg <210> 2 <211> 22 <212> DNA <213> synthetic <400> 2 22 attgggcgcc agggtggttt tt <210> 3 <211> 30 <212> DNA <213> synthetic <400> 3 30 cagggtgaag ggaagaagaa agcgaaaggt 45 <210> 4 <211> 30 <212> DNA <213> synthetic <400> 4 50 30 caggggaag ggaagaagaa agcgaaaggt 55 <210> 5 <211> 28 <212> DNA 60

<213> synthetic

	<400>	5															
	cacga	att	cc g	aggc	gatg	c tt	ccgc	tc									28
5	<210>	6															
	<211>																
10	<212>		o NA														
10				etic													
	<213>		_	etic													
15	<400>			<b>.</b>				.~~									30
	tcgac	gtc	ga c	taac	CCTT	g gc	ggaa	.agcc	;								30
	<210>	7															
20	<211>	2	3														
	<212>	D	NA														
25	<213>	s	ynth	etic	:												
to≢ 1∭	<400>	. 7															
30 30 30 30 30 30 30 30 30 30 30 30 30 3	gcatc	gcc	tc g	gaat	tcat	g gt	c										23
2000 2000 2000 2000 2000 2000	<210>	. 8															
ly Ly	<211>	. 8	36														
23	<212>	P	RT														
Control of the San	<213>	· T	herm	nus t	hern	nophi	llus										
40	<400>	. 8	3														
4 O	Met A	Asn	Ser	Glu	Ala 5	Met	Leu	Pro	Leu	Phe 10	Glu	Pro	Lys	Gly	Arg 15	Val	
45	Leu L	Leu	Val	Asp 20	Gly	His	His	Leu	Ala 25	Tyr	Arg	Thr	Phe	Phe 30	Ala	Leu	
<b>.</b>	Lys G	ly	Leu 35	Thr	Thr	Ser	Arg	Gly 40	Glu	Pro	Val	Gln	Ala 45	Val	Tyr	Gly	
50	Phe A	Ala 50	Lys	Ser	Leu	Leu	Lys 55	Ala	Leu	Lys	Glu	Asp 60	Gly	Tyr	Lys	Ala	
55	Val E 65	?he	Val	Val	Phe	Asp 70	Ala	Lys	Ala	Pro	Ser 75	Phe	Arg	His	Glu	Ala 80	
	Tyr G	3lu	Ala	Tyr	Lys 85	Ala	Gly	Arg	Ala	Pro 90	Thr	Pro	Glu	Asp	Phe 95	Pro	
60	Arg (	Gln	Leu	Ala 100	Leu	Ile	Lys	Glu	Leu 105	Val	Asp	Leu	Leu	Gly 110	Phe	Thr	

	Arg	Leu	Glu 115	Val	Pro	Gly	Tyr	Glu 120	Ala	Asp	Asp	Val	Leu 125	Ala	Thr	Leu
5	Ala	Lys 130	Lys	Ala	Glu	Lys	Glu 135	Gly	Tyr	Glu	Val	Arg 140	Ile	Leu	Thr	Ala
	Asp 145	Arg	Asp	Leu	Tyr	Gln 150	Leu	Val	Ser	Asp	Arg 155	Val	Ala	Val	Leu	His 160
10	Pro	Glu	Gly	His	Leu 165	Ile	Thr	Pro	Glu	Trp 170	Leu	Trp	Glu	Lys	Tyr 175	Gly
	Leu	Arg	Pro	Glu 180	Gln	Trp	Val	Asp	Phe 185	Arg	Ala	Leu	Val	Gly 190	Asp	Pro
15	Ser	Asp	Asn 195	Leu	Pro	Gly	Val	Lys 200	Gly	Ile	Gly	Glu	Lys 205	Thr	Ala	Leu
20	Lys	Leu 210	Leu	Lys	Glu	Trp	Gly 215	Ser	Leu	Glu	Asn	Leu 220	Leu	Lys	Asn	Leu
	Asp 225	Arg	Val	Lys	Pro	Glu 230	Asn	Val	Arg	Glu	Lys 235	Ile	Lys	Ala	His	Leu 240
25	Glu	Asp	Leu	Arg	Leu 245	Ser	Leu	Glu	Leu	Ser 250	Arg	Val	Arg	Thr	Asp 255	Leu
	Pro	Leu	Glu	Val 260	Asp	Leu	Ala	Gln	Gly 265	Arg	Glu	Pro	Asp	Arg 270	Glu	Gly
30,	Leu	Arg	Ala 275	Phe	Leu	Glu	Arg	Leu 280	Glu	Phe	Gly	Ser	Leu 285	Leu	His	Glu
35	Phe	Gly 290	Leu	Leu	Glu	Ala	Pro 295	Ala	Pro	Leu	Glu	Glu 300	Ala	Pro	Trp	Pro
Comment of the second of the s	Pro 305	Pro	Glu	Gly	Ala	Phe 310	Val	Gly	Phe	Val	Leu 315	Ser	Arg	Pro	Glu	Pro 320
40	Met	Trp	Ala	Glu	Leu 325	Lys	Ala	Leu	Ala	Ala 330	Cys	Arg	Asp	Gly	Arg 335	Val
34	His	Arg	Ala	Ala 340		Pro	Leu	Ala	Gly 345	Leu	Lys	Asp	Leu	Lys 350	Glu	Val
45	Arg	Gly	Leu 355	Leu	Ala	Lys	Asp	Leu 360	Ala	Val	Leu	Ala	Ser 365	Arg	Glu	Gly
50	Leu	Asp 370	Leu	Val	Pro	Gly	Asp 375	Asp	Pro	Met	Leu	Leu 380	Ala	Tyr	Leu	Leu
	Asp 385		Ser	Asn	Thr	Thr 390	Pro	Glu	Gly	Val	Ala 395	Arg	Arg	Tyr	Gly	Gly 400
55	Glu	Trp	Thr	Glu	Asp 405	Ala	Ala	His	Arg	Ala 410	Leu	Leu	Ser	Glu	Arg 415	Leu
60	His	Arg	Asn	Leu 420		Lys	Arg	Leu	Glu 425		Glu	Glu	Lys	Leu 430	Leu	Trp
o o	Leu	Tyr	His 435		Val	Glu	Lys	Pro 440		Ser	Arg	Val	Leu 445	Ala	His	Met

	Glu	Ala 450	Thr	Gly	Val	Arg	Arg 455	Asp	Val	Ala	Tyr	Leu 460	Gln	Ala	Leu	Ser
5	Leu 465	Glu	Leu	Ala	Glu	Glu 470	Ile	Arg	Arg	Leu	Glu 475	Glu	Glu	Val	Phe	Arg 480
	Leu	Ala	Gly	His	Pro 485	Phe	Asn	Leu	Asn	Ser 490	Arg	Asp	Gln	Leu	Glu 495	Arg
10	Val	Leu	Phe	Asp 500	Glu	Leu	Arg	Leu	Pro 505	Ala	Leu	Gly	Lys	Thr 510	Gln	Lys
a.e.	Thr	Gly	Lys 515	Arg	Ser	Thr	Ser	Ala 520	Ala	Val	Leu	Glu	Ala 525	Leu	Arg	Glu
15	Ala	His 530	Pro	Ile	Val	Glu	Lys 535	Ile	Leu	Gln	His	Arg 540	Glu	Leu	Thr	Lys
20	Leu 545	Lys	Asn	Thr	Tyr	Val 550	Asp	Pro	Leu	Pro	Ser 555	Leu	Val	His	Pro	Arg 560
	Thr	Gly	Arg	Leu	His 565	Thr	Arg	Phe	Asn	Gln 570	Thr	Ala	Thr	Ala	Thr 575	Gly
25	Arg	Leu	Ser	Ser 580	Ser	Asp	Pro	Asn	Leu 585	Gln	Asn	Ile	Pro	Val 590	Arg	Thr
Share Marie Marie	Pro	Leu	Gly 595	Gln	Arg	Ile	Arg	Arg 600	Ala	Phe	Val	Ala	Glu 605	Ala	Gly	Trp
30	Ala	Leu 610	Val	Ala	Leu	Asp	Tyr 615	Ser	Gln	Ile	Glu	Leu 620	Arg	Val	Leu	Ala
3.5 miles	His 625	Leu	Ser	Gly	Asp	Glu 630	Asn	Leu	Ile	Arg	Val 635	Phe	Gln	Glu	Gly	Lys 640
Such Such	Asp	Ile	His	Thr	Gln 645	Thr	Ala	Ser	Trp	Met 650	Phe	Gly	Val	Pro	Pro 655	Glu
40	Ala	Val	Asp	Pro 660	Leu	Met	Arg	Arg	Ala 665	Ala	Lys	Thr	Val	Asn 670	Phe	Gly
}=4	Val		Tyr 675		Met				Arg				Glu 685	Leu	Ala	Ile
45	Pro	Tyr 690	Glu	Glu	Ala	Val	Ala 695	Phe	Ile	Glu	Arg	Tyr 700	Phe	Gln	Ser	Phe
50	Pro 705		Val	Arg	Ala	Trp 710		Glu	Lys	Thr	Leu 715	Glu	Glu	Gly	Arg	Lys 720
	Arg	Gly	Tyr	Val	Glu 725	Thr	Leu	Phe	Gly	Arg 730	Arg	Arg	Tyr	Val	Pro 735	Asp
55	Leu	Asn	Ala	Arg 740		Lys	Ser	Val	Arg 745	Glu	Ala	Ala	Glu	Arg 750	Met	Ala
60	Phe	Asn	Met 755		Val	Gln	. Gly	Thr 760	Ala	Ala	Asp	Leu	Met 765	Lys	Leu	Ala
50	Met	Val 770		Leu	Phe	Pro	Arg 775		Arg	Glu	. Met	Gly 780	Ala	Arg	Met	Leu

	Leu Gln 785	Val	His	Asp	Glu 790	Leu	Leu	Leu	Glu	Ala 795	Pro	Gln	Ala	Arg	Ala 800	
5	Glu Glu	Val	Ala	Ala 805	Leu	Ala	Lys	Glu	Ala 810	Met	Glu	Lys	Ala	Tyr 815	Pro	
	Leu Ala	Val	Pro 820	Leu	Glu	Val	Glu	Val 825	Gly	Met	Gly	Glu	Asp 830	Trp	Leu	
10	Ser Ala	Lys 835														
	<210>	9														
15	<211>	2511														
	<212>	DNA														
20	<213>	Ther	mus	therm	noph:	ilus										
	<400>	9														
25	atgaatt	ccg	aggc	gatgo	ct to	ccgct	cttt	ga	accca	aaag	gcc	gggt	cct	cctg	gtggac	60
25	ggccacc	acc	tggc	ctaco	eg e	accti	ctto	c gc	cctg	aagg	gcc	tcac	cac	gagc	cggggc	120
	gaaccgg	tgc	aggc	ggtct	ta c	ggcti	cgc	aa	gagc	ctcc	tca	aggc	cct	gaag	gaggad	180
30	gggtaca	agg	ccgt	cttc	gt g	gtcti	tgad	gc.	caag	gccc	cct	cctt	ccg	ccac	gaggco	240
300	tacgagg	cct	acaa	ggcg	aa a	aggg	caca	g ac	cccc	gagg	act	tccc	ccg	gcag	ctcgcc	300
	ctcatca	agg	agct	ggtg	ga c	ctcc'	tgggg	g tt	tacc	cgcc	tcg	aggt	ccc	cggc	tacgag	360
35	gcggacg	acg	ttct	cgcca	ac c	ctgg	ccaa	g aa	ggcg	gaaa	agg.	aggg	gta	cgag	gtgcgc	420
A conf	atcctca	ccg	ccga	ccgc	ga c	ctct	accaa	a ct	cgtc	tccg	acc	gcgt	cgc	cgtc	ctccac	480
40	cccgagg	gcc	acct	catc	ac c	ccgg	agtg	g ct	ttgg	gaga	agt	acgg	cct	cagg	ccggag	540
500 f 1000 1000 1000	cagtggg	tgg	actt	ccgc	gc c	ctcg	tggg	g ga	cccc	tccg	aca	acct	ccc	cggg	gtcaag	600
	ggcatcg	ggg	agaa	gacc	gc c	ctca	agct	c ct	caag	gagt	999	gaag	cct	ggaa	aaccto	: 660
45	ctcaaga	.acc	tgga	ccgg	gt a	aagc	caga	a aa	cgtc	cggg	aga	agat	caa	ggcc	caccto	720
	gaagacc	tca	ggct	ctcc	tt g	gagc	tctc	c cg	ggtg	cgca	ccg	acct	ccc	cctg	gaggto	780
50	gacctcg	ccc	aggg	gcgg	ga g	cccg	accg	g ga	9999	ctta	<b>a</b> aa	cctt	cct	ggag	aggcts	840
	gagttcg	gca	gcct	cctc	ca c	gagt	tcgg	c ct	cctg	gagg	ccc	ccgc	ccc	cctg	gaggag	900
	gccccct	.ggc	cccc	gccg	ga a	.gggg	cctt	c gt	gggc	ttcg	tcc	tctc	ccg	cccc	gagcc	960
55	atgtggg	ıcgg	agct	taaa	gc c	ctgg	ccgc	c tg	cagg	gacg	gcc	gggt	gca	ccgg	gcagca	a 1020
	gacccct	tgg	cggg	gcta	aa g	gacc	tcaa	g ga	ggtc	cggg	gcc	tcct	cgc	caag	gaccto	1080
60	gccgtct	tgg	cctc	gagg	ga g	gggc	taga	c ct	cgtg	cccg	999	acga	CCC	catg	ctcct	1140
	gcctaco	ctcc	tgga	cccc	tc c	aaca	ccac	c cc	cgag	9999	tgg	cgcg	gcg	ctac	999999	1200

	gagtggacgg	aggacgccgc	ccaccgggcc	ctcctctcgg	agaggeteea	teggaacete	1200
	cttaagcgcc	tcgagggga	ggagaagctc	ctttggctct	accacgaggt	ggaaaagccc	1320
5	ctctcccggg	tcctggccca	catggaggcc	accggggtac	ggcgggacgt	ggcctacctt	1380
	caggcccttt	ccctggagct	tgcggaggag	atccgccgcc	tcgaggagga	ggtcttccgc	1440
	ttggcgggcc	accccttcaa	cctcaactcc	cgggaccagc	tggaaagggt	gctctttgac	1500
LO	gagcttaggc	ttcccgcctt	ggggaagacg	caaaagacag	gcaagcgctc	caccagcgcc	1560
	gcggtgctgg	g aggccctacg	ggaggcccac	cccatcgtgg	agaagatcct	ccagcaccgg	1620
15	gagctcacca	agctcaagaa	cacctacgtg	gaccccctcc	caagcctcgt	ccacccgagg	1680
	acgggccgcc	tccacacccg	cttcaaccag	acggccacgg	ccacggggag	gcttagtagc	1740
	tccgacccca	acctgcagaa	cateceegte	cgcaccccct	tgggccagag	gatccgccgg	1800
20	gccttcgtgg	g ccgaggcggg	ttgggcgttg	gtggccctgg	actatagcca	gatagagctc	1860
	cgcgtcctcg	g cccacctctc	cggggacgaa	aacctgatca	gggtcttcca	ggagggaag	1920
25	gacatccaca	a cccagaccgc	aagctggatg	ttcggcgtcc	ccccggaggc	cgtggacccc	1980
3	ctgatgcgc	gggcggccaa	gacggtgaac	ttcggcgtcc	tctacggcat	gtccgcccat	2040
	aggetetee	aggagettge	catcccctac	gaggaggcgg	tggcctttat	agagcgctac	2100
30 ]	ttccaaagct	tccccaaggt	gcgggcctgg	atagaaaaga	ccctggagga	ggggaggaag	2160
	cggggctacg	g tggaaaccct	cttcggaaga	aggcgctacg	tgcccgacct	caacgcccgg	2220
35	gtgaagagcg	g tcagggaggc	cgcggagcgc	atggccttca	acatgcccgt	ccagggcacc	2280
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	gccgccgac	c tcatgaagct	cgccatggtg	aagctcttcc	cccgcctccg	ggagatgggg	2340
40	gcccgcatgo	c teetecaggt	ccacgacgag	ctcctcctgg	aggcccccca	agcgcgggcc	2400
40	gaggaggtgg	g cggctttggc	caaggaggcc	atggagaagg	cctatcccct	cgccgtgccc	2460
	ctggaggtg	g aggtggggat	gggggaggac	tggctttccg	ccaagggtta	g	2511
45	<210> 10						
	<211> 26						
50	<212> DN	A					
	<213> sy	nthetic					
55	<400> 10						
<b>9</b> 9	caggaggag	c tcgttgtgga	cctgga				26
60	<210> 11						
60	<211> 83	6					

<212> PRT Thermus thermophilus 5 <400> 11 Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu 10 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly 15 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala 20 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr 25 3 3 3 3 4 105 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly 165 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu 45 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu 215 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu 50 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly 55

300

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu

Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro

	Pro 305	Pro	Glu	Gly	Ala	Phe 310	Val	Gly	Phe	Val	Leu 315	Ser	Arg	Pro	Glu	Pro 320
5	Met	Trp	Ala	Glu	Leu 325	Lys	Ala	Leu	Ala	Ala 330	Cys	Arg	Asp	Gly	Arg 335	Val
	His	Arg	Ala	Ala 340	Asp	Pro	Leu	Ala	Gly 345	Leu	Lys	Asp	Leu	Lys 350	Glu	Val
10	Arg	Gly	Leu 355	Leu	Ala	Lys	Asp	Leu 360	Ala	Val	Leu	Ala	Ser 365	Arg	Glu	Gly
1 5	Leu	Asp 370	Leu	Val	Pro	Gly	Asp 375	Asp	Pro	Met	Leu	Leu 380	Ala	Tyr	Leu	Leu
15	Asp 385	Pro	Ser	Asn	Thr	Thr 390	Pro	Glu	Gly	Val	Ala 395	Arg	Arg	Tyr	Gly	Gly 400
20	Glu	Trp	Thr	Glu	Asp 405	Ala	Ala	His	Arg	Ala 410	Leu	Leu	Ser	Glu	Arg 415	Leu
	His	Arg	Asn	Leu 420	Leu	Lys	Arg	Leu	Glu 425	Gly	Glu	Glu	Lys	Leu 430	Leu	Trp
25	Leu	Tyr	His 435	Glu	Val	Glu	Lys	Pro 440	Leu	Ser	Arg	Val	Leu 445	Ala	His	Met
3.0 	Glu	Ala 450	Thr	Gly	Val	Arg	Arg 455	Asp	Val	Ala	Tyr	Leu 460	Gln	Ala	Leu	Ser
30	Leu 465	Glu	Leu	Ala	Glu	Glu 470	Ile	Arg	Arg	Leu	Glu 475	Glu	Glu	Val	Phe	Arg 480
35	Leu	Ala	Gly	His	Pro 485	Phe	Asn	Leu	Asn	Ser 490	Arg	Asp	Gln	Leu	Glu 495	Arg
Action States	Val	Leu	Phe	Asp 500	Glu	Leu	Arg	Leu	Pro 505	Ala	Leu	Gly	Lys	Thr 510	Gln	Lys
40	Thr	Gly	Lys 515	Arg	Ser	Thr	Ser	Ala 520	Ala	Val	Leu	Glu	Ala 525	Leu	Arg	Glu
45	Ala	His 530	Pro	Ile	Val	Glu	Lys 535	Ile	Leu	Gln	His	Arg 540	Glu	Leu	Thr	Lys
13	Leu 545	Lys	Asn	Thr	Tyr	Val 550	Asp	Pro	Leu	Pro	Ser 555	Leu	Val	His	Pro	Arg 560
50	Thr	Gly	Arg	Leu	His 565	Thr	Arg	Phe	Asn	Gln 570	Thr	Ala	Thr	Ala	Thr 575	Gly
	Arg	Leu	Ser	Ser 580	Ser	Asp	Pro	Asn	Leu 585	Gln	Asn	Ile	Pro	Val 590	Arg	Thr
55	Pro	Leu	Gly 595	Gln	Arg	Ile	Arg	Arg 600	Ala	Phe	۷al	Ala	Glu 605	Ala	Gly	Trp
60	Ala	Leu 610	Val	Ala	Leu	Asp	Tyr 615		Gln	Ile	Glu	Leu 620	Arg	Val	Leu	Ala
	His 625		Ser	Gly	Asp	Glu 630		Leu	Ile	Arg	Val 635	Phe	Gln	Glu	Gly	Lys 640

	Asp	Ile	His	Thr	Gln 645	Thr	Ala	Ser	Trp	Met 650	Phe	Gly	Val	Pro	Pro 655	Glu	
5	Ala	Val	Asp	Pro 660	Leu	Met	Arg	Arg	Ala 665	Ala	Lys	Thr	Val	Asn 670	Phe	Gly	
	Val	Leu	Tyr 675	Gly	Met	Ser	Ala	His 680	Arg	Leu	Ser	Gln	Glu 685	Leu	Ala	Ile	
10	Pro	Tyr 690	Glu	Glu	Ala	Val	Ala 695	Phe	Ile	Glu	Arg	Tyr 700	Phe	Gln	Ser	Phe	
. r	Pro 705	Lys	Val	Arg	Ala	Trp 710	Ile	Glu	Lys	Thr	Leu 715	Glu	Glu	Gly	Arg	Lys 720	
15	Arg	Gly	Tyr	Val	Glu 725	Thr	Leu	Phe	Gly	Arg 730	Arg	Arg	Tyr	Val	Pro 735	Asp	
20	Leu	Asn	Ala	Arg 740	Val	Lys	Ser	Val	Arg 745	Glu	Ala	Ala	Glu	Arg 750	Met	Ala	
	Phe	Asn	Met 755	Pro	Val	Gln	Gly	Thr 760	Ala	Ala	Asp	Leu	Met 765	Lys	Leu	Ala	
25	Met	Val 770	Lys	Leu	Phe	Pro	Arg 775	Leu	Arg	Glu	Met	Gly 780	Ala	Arg	Met	Leu	
25	Leu 785	Gln	Val	His	Asn	Glu 790	Leu	Leu	Leu	Glu	Ala 795	Pro	Gln	Ala	Arg	Ala 800	
30	Glu	Glu	Val	Ala	Ala 805	Leu	Ala	Lys	Glu	Ala 810	Met	Glu	Lys	Ala	Tyr 815	Pro	
35	Leu	Ala	Val	Pro 820	Leu	Glu	Val	Glu	Val 825	Gly	Met	Gly	Glu	Asp 830	Trp	Leu	
Access of the second se	Ser	Ala	Lys 835	Gly													
40	<21	0 >	12														
	<21	1>	2511														
45	<21	2>	DNA														
	<21	3>	Ther	mus	ther	moph	ilus										
	<40	0>	12														
50	atg	aatt	ccg	aggc	gatg	ct t	ccgc	tctt	t ga	accc	aaag	gcc	gggt	cct	cctg	gtggac	60
	ggc	cacc	acc	tggc	ctac	cg c	acct	tctt	c gc	cctg	aagg	gcc	tcac	cac	gagc	cggggc	120
55	gaa	ccgg	tgc	aggc	ggta	ta c	ggct	tcgc	c aa	gagc	ctcc	tca	aggc	cct	gaag	gaggac	180
	999	taca	.agg	ccgt	cttc	gt g	gtct	ttga	.c gc	caag	gccc	cct	cctt	ccg	ccac	gaggcc	240
	tac	gagg	cct	acaa	ggcg		aggg	cccc	g ac	cccc	gagg	act	tccc	ccg	gcag	ctcgcc	300
60	ctc	atca	agg	agct	ggtg	ga c	ctcc	tggg	g tt	tacc	cgcc	tc9	aggt	ccc	cggc	tacgag	360

Ç	gcggacgacg	ttctcgccac	cctggccaag	aaggcggaaa	aggaggggta	cgaggtgcgc	420
ä	atcctcaccg	ccgaccgcga	cctctaccaa	ctcgtctccg	accgcgtcgc	cgtcctccac	480
Ó	cccgagggcc	acctcatcac	cccggagtgg	ctttgggaga	agtacggcct	caggccggag	540
(	cagtgggtgg	acttccgcgc	cctcgtgggg	gacccctccg	acaacctccc	cggggtcaag	600
9	ggcatcgggg	agaagaccgc	cctcaagctc	ctcaaggagt	ggggaagcct	ggaaaacctc	660
	ctcaagaacc	tggaccgggt	aaagccagaa	aacgtccggg	agaagatcaa	ggcccacctg	720
9	gaagacctca	ggctctcctt	ggagctctcc	cgggtgcgca	ccgacctccc	cctggaggtg	780
9	gacctcgccc	aggggcggga	gcccgaccgg	gaggggctta	gggccttcct	ggagaggctg	840
9	gagttcggca	gcctcctcca	cgagttcggc	ctcctggagg	ccccgcccc	cctggaggag	900
!	gccccctggc	ccccgccgga	aggggccttc	gtgggcttcg	tcctctcccg	ccccgagccc	960
,	atgtgggcgg	agcttaaagc	cctggccgcc	tgcagggacg	gccgggtgca	ccgggcagca	1020
	gaccccttgg	cggggctaaa	ggacctcaag	gaggtccggg	gcctcctcgc	caaggacctc	1080
,	gccgtcttgg	cctcgaggga	ggggctagac	ctcgtgcccg	gggacgaccc	catgctcctc	1140
,	gcctacctcc	tggacccctc	caacaccacc	cccgaggggg	tggcgcggcg	ctacgggggg	1200
	gagtggacgg	aggacgccgc	ccaccgggcc	ctcctctcgg	agaggctcca	teggaacete	1260
	cttaagcgcc	tcgaggggga	ggagaagctc	ctttggctct	accacgaggt	ggaaaagccc	1320
	ctctcccggg	tcctggccca	catggaggcc	accggggtac	ggcgggacgt	ggcctacctt	1380
	caggcccttt	ccctggagct	tgcggaggag	atccgccgcc	tcgaggagga	ggtcttccgc	1440
	ttggcgggcc	accccttcaa	cctcaactcc	cgggaccagc	tggaaagggt	gctctttgac	1500
	gagcttaggc	ttcccgcctt	ggggaagacg	caaaagacag	gcaagcgctc	caccagcgcc	1560
	gcggtgctgg	aggccctacg	ggaggcccac	cccatcgtgg	agaagatcct	ccagcaccgg	1620
	gagctcacca	agctcaagaa	cacctacgtg	gaccccctcc	caagcctcgt	ccacccgagg	1680
	acgggccgcc	tccacacccg	cttcaaccag	acggccacgg	ccacggggag	gcttagtagc	1740
	tccgacccca	acctgcagaa	catccccgtc	cgcaccccct	tgggccagag	gatccgccgg	1800
	gccttcgtgg	ccgaggcggg	ttgggcgttg	gtggccctgg	actatagcca	gatagagctc	1860
	cgcgtcctcg	cccacctctc	cggggacgaa	aacctgatca	gggtcttcca	ggagggaag	1920
	gacatccaca	cccagaccgc	aagctggatg	ttcggcgtcc	ccccggaggc	cgtggacccc	1980
	ctgatgcgcc	gggcggccaa	gacggtgaac	ttcggcgtcc	tctacggcat	gtccgcccat	2040
	aggctctccc	aggagcttgc	catcccctac	gaggaggcgg	tggcctttat	agagcgctac	2100
	ttccaaagct	tccccaaggt	gcgggcctgg	atagaaaaga	ccctggagga	ggggaggaag	2160
	cggggctacg	tggaaaccct	cttcggaaga	aggcgctacg	tgcccgacct	caacgcccgg	2220

	gtgaagagcg	tcagggaggc	cgcggagcgc	atggccttca	acatgcccgt	ccagggcacc	2280
	gccgccgacc	tcatgaagct	cgccatggtg	aagctcttcc	cccgcctccg	ggagatgggg	2340
5	gcccgcatgc	tcctccaggt	ccacaacgag	ctcctcctgg	aggcccccca	agcgcgggcc	2400
	gaggaggtgg	cggctttggc	caaggaggcc	atggagaagg	cctatcccct	cgccgtgccc	2460
10	ctggaggtgg	aggtggggat	gggggaggac	tggctttccg	ccaagggtta	g	2511
	<210> 13						
1 5	<211> 58						
15	<212> DNA						
	<213> synt	hetic					
20	<400> 13						
	tgcctgcagg	tcgacgctag	ctagtggtgg	tggtggtggt	gacccttggc	ggaaagcc	58
25	<210> 14						
30,744,75	<211> 2526	,					
	<212> DNA						
3 U	<213> Ther	mus thermop	philus				
	<400> 14						
35	atgaattccg	aggcgatgct	teegetettt	gaacccaaag	gccgggtcct	cctggtggac	60
1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	ggccaccacc	tggcctaccg	caccttcttc	gccctgaagg	gcctcaccac	gagccggggc	120
400	gaaccggtgc	aggcggtcta	cggcttcgcc	aagagcctcc	tcaaggccct	gaaggaggac	180
4 O constraint	gggtacaagg	ccgtcttcgt	ggtctttgac	gccaaggccc	cctccttccg	ccacgaggcc	240
	tacgaggcct	acaaggcggg	gagggccccg	acccccgagg	acttcccccg	gcagctcgcc	300
45	ctcatcaagg	agctggtgga	cctcctgggg	tttacccgcc	tcgaggtccc	cggctacgag	360
	gcggacgacg	ttctcgccac	cctggccaag	aaggcggaaa	aggaggggta	cgaggtgcgc	420
50	atcctcaccg	ccgaccgcga	cctctaccaa	ctcgtctccg	accgcgtcgc	cgtcctccac	480
50	cccgagggcc	acctcatcac	cccggagtgg	ctttgggaga	agtacggcct	caggccggag	540
	cagtgggtgg	acttccgcgc	cctcgtgggg	gacccctccg	acaacctccc	cggggtcaag	600
55	ggcatcgggg	agaagaccgc	cctcaagctc	ctcaaggagt	ggggaagcct	ggaaaacctc	660
	ctcaagaacc	tggaccgggt	aaagccagaa	aacgtccggg	agaagatcaa	ggcccacctg	720
60	gaagacctca	ggctctcctt	ggagctctcc	cgggtgcgca	ccgacctccc	cctggaggtg	780
00	gacctcgccc	aggggcggga	gcccgaccgg	gaggggctta	gggccttcct	ggagaggctg	840

	gagttcggca	gcctcctcca	cgagttcggc	ctcctggagg	ccccgcccc	cctggaggag	900
	gccccctggc	ccccgccgga	aggggccttc	gtgggcttcg	tcctctcccg	ccccgagccc	960
5	atgtgggcgg	agcttaaagc	cctggccgcc	tgcagggacg	gccgggtgca	ccgggcagca	1020
	gaccccttgg	cggggctaaa	ggacctcaag	gaggtccggg	gcctcctcgc	caaggacctc	1080
1.0	gccgtcttgg	cctcgaggga	ggggctagac	ctcgtgcccg	gggacgaccc	catgctcctc	1140
10	gcctacctcc	tggacccctc	caacaccacc	cccgaggggg	tggcgcggcg	ctacgggggg	1200
	gagtggacgg	aggacgccgc	ccaccgggcc	ctcctctcgg	agaggctcca	tcggaacctc	1260
15	cttaagcgcc	tcgaggggga	ggagaagctc	ctttggctct	accacgaggt	ggaaaagccc	1320
	ctctcccggg	tcctggccca	catggaggcc	accggggtac	ggcgggacgt	ggcctacctt	1380
2.0	caggcccttt	ccctggagct	tgcggaggag	atccgccgcc	tcgaggagga	ggtcttccgc	1440
20	ttggcgggcc	accccttcaa	cctcaactcc	cgggaccagc	tggaaagggt	gctctttgac	1500
	gagcttaggc	ttcccgcctt	ggggaagacg	caaaagacag	gcaagcgctc	caccagcgcc	1560
25	gcggtgctgg	aggccctacg	ggaggcccac	cccatcgtgg	agaagatcct	ccagcaccgg	1620
	gagctcacca	agctcaagaa	cacctacgtg	gaccccctcc	caagcctcgt	ccacccgagg	1680
	acgggccgcc	tccacacccg	cttcaaccag	acggccacgg	ccacggggag	gcttagtagc	1740
30	tccgacccca	acctgcagaa	catccccgtc	cgcaccccct	tgggccagag	gatccgccgg	1800
	gccttcgtgg	ccgaggcggg	ttgggcgttg	gtggccctgg	actatagcca	gatagagctc	1860
35	cgcgtcctcg	cccacctctc	cggggacgaa	aacctgatca	gggtcttcca	ggagggaag	1920
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	gacatccaca	cccagaccgc	aagctggatg	ttcggcgtcc	ccccggaggc	cgtggacccc	1980
40	ctgatgcgcc	gggcggccaa	gacggtgaac	ttcggcgtcc	tctacggcat	gtccgcccat	2040
40	aggctctccc	aggagcttgc	catcccctac	gaggaggcgg	tggcctttat	agagcgctac	2100
	ttccaaagct	tccccaaggt	gcgggcctgg	atagaaaaga	ccctggagga	ggggaggaag	2160
45	cggggctacg	tggaaaccct	cttcggaaga	aggcgctacg	tgcccgacct	caacgcccgg	2220
	gtgaagagcg	tcagggaggc	cgcggagcgc	atggccttca	acatgcccgt	ccagggcacc	2280
50	gccgccgacc	tcatgaagct	cgccatggtg	aagctcttcc	cccgcctccg	ggagatgggg	2340
50	gcccgcatgc	tcctccaggt	ccacaacgag	ctcctcctgg	aggcccccca	agegegggee	2400
	gaggaggtgg	cggctttggc	caaggaggcc	atggagaagg	cctatcccct	cgccgtgccc	2460
55	ctggaggtgg	aggtggggat	gggggaggac	tggctttccg	ccaagggtca	ccaccaccac	2520
	caccac						2526

60 <210> 15

<211> 842

<212> PRT Thermus thermophilus <400> 15 5 Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu 10 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly 15 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala 20 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr 25 30 35 40 11 1 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His 155 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly 165 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro 185 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu 45 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu 50 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly 55 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu 60 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro

295

	Pro 305	Pro	Glu	Gly	Ala	Phe 310	Val	Gly	Phe	Val	Leu 315	Ser	Arg	Pro	Glu	Pro 320
5	Met	Trp	Ala	Glu	Leu 325	Lys	Ala	Leu	Ala	Ala 330	Cys	Arg	Asp	Gly	Arg 335	Val
	His	Arg	Ala	Ala 340	Asp	Pro	Leu	Ala	Gly 345	Leu	Lys	Asp	Leu	Lys 350	Glu	Val
10	Arg	Gly	Leu 355	Leu	Ala	Lys	Asp	Leu 360	Ala	Val	Leu	Ala	Ser 365	Arg	Glu	Gly
15	Leu	Asp 370	Leu	Val	Pro	Gly	Asp 375	Asp	Pro	Met	Leu	Leu 380	Ala	Tyr	Leu	Leu
15	Asp 385	Pro	Ser	Asn	Thr	Thr 390	Pro	Glu	Gly	Val	Ala 395	Arg	Arg	Tyr	Gly	Gly 400
20	Glu	Trp	Thr	Glu	Asp 405	Ala	Ala	His	Arg	Ala 410	Leu	Leu	Ser	Glu	Arg 415	Leu
	His	Arg	Asn	Leu 420	Leu	Lys	Arg	Leu	Glu 425	Gly	Glu	Glu	Lys	Leu 430	Leu	Trp
25	Leu	Tyr	His 435	Glu	Val	Glu	Lys	Pro 440	Leu	Ser	Arg	Val	Leu 445	Ala	His	Met
dien für für	Glu	Ala 450	Thr	Gly	Val	Arg	Arg 455	Asp	Val	Ala	Tyr	Leu 460	Gln	Ala	Leu	Ser
3 O	Leu 465	Glu	Leu	Ala	Glu	Glu 470	Ile	Arg	Arg	Leu	Glu 475	Glu	Glu	Val	Phe	Arg 480
35	Leu	Ala	Gly	His	Pro 485	Phe	Asn	Leu	Asn	Ser 490	Arg	Asp	Gln	Leu	Glu 495	Arg
	Val	Leu	Phe	Asp 500	Glu	Leu	Arg	Leu	Pro 505	Ala	Leu	Gly	Lys	Thr 510	Gln	Lys
40	Thr	Gly	Lys 515	Arg	Ser	Thr	Ser	Ala 520	Ala	Val	Leu	Glu	Ala 525	Leu	Arg	Glu
45	Ala	His 530	Pro	Ile	Val	Glu	Lys 535	Ile	Leu	Gln	His	Arg 540	Glu	Leu	Thr	Lys
43	Leu 545	Lys	Asn	Thr	Tyr	Val 550	Asp	Pro	Leu	Pro	Ser 555	Leu	Val	His	Pro	Arg 560
50	Thr	Gly	Arg	Leu	His 565	Thr	Arg	Phe	Asn	Gln 570	Thr	Ala	Thr	Ala	Thr 575	Gly
	Arg	Leu	Ser	Ser 580	Ser	Asp	Pro	Asn	Leu 585		Asn	Ile	Pro	Val 590	Arg	Thr
55	Pro	Leu	Gly 595		Arg	Ile	Arg	Arg 600		Phe	Val	Ala	Glu 605	Ala	Gly	Trp
60		610					615					620			Leu	
• •	His 625		Ser	Gly	Asp	Glu 630		Leu	Ile	Arg	Val 635	Phe	Gln	Glu	Gly	Lys 640

	Asp	Ile	His	Thr	Gln 645	Thr	Ala	Ser	Trp	Met 650	Phe	Gly	Val	Pro	Pro 655	Glu	
5	Ala	Val	Asp	Pro 660	Leu	Met	Arg	Arg	Ala 665	Ala	Lys	Thr	Val	Asn 670	Phe	Gly	
	Val	Leu	Tyr 675	Gly	Met	Ser	Ala	His 680	Arg	Leu	Ser	Gln	Glu 685	Leu	Ala	Ile	
10	Pro	Tyr 690	Glu	Glu	Ala	Val	Ala 695	Phe	Ile	Glu	Arg	Tyr 700	Phe	Gln	Ser	Phe	
15	Pro 705	Lys	Val	Arg	Ala	Trp 710	Ile	Glu	Lys	Thr	Leu 715	Glu	Glu	Gly	Arg	Lys 720	
15	Arg	Gly	Tyr	Val	Glu 725	Thr	Leu	Phe	Gly	Arg 730	Arg	Arg	Tyr	Val	Pro 735	Asp	
20	Leu	Asn	Ala	Arg 740	Val	Lys	Ser	Val	Arg 745	Glu	Ala	Ala	Glu	Arg 750	Met	Ala	
	Phe	Asn	Met 755	Pro	Val	Gln	Gly	Thr 760	Ala	Ala	Asp	Leu	Met 765	Lys	Leu	Ala	
25	Met	Val 770	Lys	Leu	Phe	Pro	Arg 775	Leu	Arg	Glu	Met	Gly 780	Ala	Arg	Met	Leu	
	Leu 785	Gln	Val	His	Asn	Glu 790	Leu	Leu	Leu	Glu	Ala 795	Pro	Gln	Ala	Arg	Ala 800	
25	Glu	Glu	Val	Ala	Ala 805	Leu	Ala	Lys	Glu	Ala 810	Met	Glu	Lys	Ala	Tyr 815	Pro	
35	Leu	Ala	Val	Pro 820	Leu	Glu	Val	Glu	Val 825	Gly	Met	Gly	Glu	Asp 830	Trp	Leu	
4 Q	Ser	Ala	Lys 835	Gly	His	His	His	His 840	His	His							
40	<21	0 >	16														
	<21	1>	31														
45	<21	2>	DNA														
	<21	3 >	synt	heti	С												
	<40	0>	16														
50	gcc	tgca	ggg	gcgg	ccgc	gt g	cacc	9999	са								31
	<21	0>	17														
55	<21	1>	26														
	<21	2>	DNA														
60	<21	3 >	synt	heti	.C												

	ctcctgga	cc cttcgaacac	cacccc				26
	<210> 1	8					
5	<211> 2	3					
	<212> D	NA					
10	<213> s	ynthetic					
	<400> 1	8					
15	gtcctggc	cc atatggaggc	cac				23
	<210> 1	9					
2.0	<211> 2	526					
20	<212> D	NA					
	<213> T	hermus thermo	philus				
25	<400> 1	9					
25 May 100 May	atgaattc	cg aggcgatgct	tccgctcttt	gaacccaaag	gccgggtcct	cctggtggac	60
	ggccacca	cc tggcctaccg	g caccttcttc	gccctgaagg	gcctcaccac	gagccggggc	120
30	gaaccggt	gc aggcggtcta	a cggcttcgcc	aagagcctcc	tcaaggccct	gaaggaggac	180
	gggtacaa	.gg ccgtcttcgt	ggtctttgac	gccaaggccc	cctccttccg	ccacgaggcc	240
35	tacgaggc	ct acaaggcggg	g gagggccccg	acccccgagg	acttcccccg	gcagctcgcc	300
	ctcatcaa	.gg agctggtgga	a cctcctgggg	tttacccgcc	tcgaggtccc	cggctacgag	360
	gcggacga	cg ttctcgcca	c cctggccaag	aaggcggaaa	aggaggggta	cgaggtgcgc	420
40	atcctcac	cg ccgaccgcg	a cctctaccaa	ctcgtctccg	accgcgtcgc	cgtcctccac	480
	cccgaggg	cc acctcatca	c cccggagtgg	ctttgggaga	agtacggcct	caggccggag	540
45	cagtgggt	gg acttccgcg	c cctcgtgggg	gacccctccg	acaacctccc	cggggtcaag	600
	ggcatcgg	gg agaagaccg	c cctcaagctc	ctcaaggagt	ggggaagcct	ggaaaacctc	660
	ctcaagaa	cc tggaccggg	t aaagccagaa	aacgtccggg	agaagatcaa	ggcccacctg	720
50	gaagacct	ca ggctctcct	t ggagctctcc	cgggtgcgca	ccgacctccc	cctggaggtg	780
	gacctcgc	cc aggggcggg	a gcccgaccgg	gaggggctta	gggccttcct	ggagaggctg	840
55	gagttcgg	gca gcctcctcc	a cgagttcggc	ctcctggagg	ccccgcccc	cctggaggag	900
	gccccctg	ge cecegeegg	a aggggccttc	gtgggcttcg	tcctctcccg	ccccgagccc	960
	atgtggg	gg agcttaaag	c cctggccgcc	tgcaggggcg	gccgcgtgca	ccgggcagca	1020
60	gacccctt	gg cggggctaa	a ggacctcaag	gaggtccggg	gcctcctcgc	caaggacctc	1080

```
gccgtcttgg cctcgaggga ggggctagac ctcgtgcccg gggacgaccc catgctcctc
                                                                            1200
      gectacetee tggaceette gaacaceace eeegaggggg tggegeggeg etacgggggg
                                                                            1260
      gagtggacgg aggacgccgc ccaccgggcc ctcctctcgg agaggctcca tcggaacctc
5
                                                                            1320
      cttaagcgcc tcgagggga ggagaagctc ctttggctct accacgaggt ggaaaagccc
                                                                            1380
      ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt
10
       caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc
                                                                            1440
                                                                            1500
       ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac
                                                                            1560
      gagettagge ttecegeett ggggaagaeg caaaagaeag geaagegete caecagegee
15
       gcggtgctgg aggccctacg ggaggcccac cccatcgtgg agaagatcct ccagcaccgg
                                                                            1620
       gageteacea ageteaagaa eacetaegtg gaceceetee caageetegt ecaceegagg
                                                                            1680
20
                                                                            1740
       acgggccgcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc
       tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg
                                                                            1800
                                                                            1860
       gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc
1920
       cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggtcttcca ggaggggaag
                                                                            1980
       gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccggaggc cgtggacccc
       ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccat
                                                                            2040
                                                                            2100
       aggetetece aggagettge catecectae gaggaggegg tggeetttat agagegetae
                                                                            2160
       ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag
                                                                            2220
       cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg
                                                                            2280
       gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc
       gccgccgacc tcatgaagct cgccatggtg aagctettee cccgcctccg ggagatgggg
                                                                            2340
       gcccgcatgc tcctccaggt ccacaacgag ctcctcctgg aggcccccca agcgcgggcc
                                                                            2400
                                                                            2460
       gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc
45
       ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac
                                                                            2520
                                                                            2526
       caccac
50
       <210>
              20
       <211>
              842
55
       <212>
              PRT
       <213>
              Thermus thermophilus
```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val

<400>

60

	1				5					10					15	
	Leu	Leu	Val	Asp 20	Gly	His	His	Leu	Ala 25	Tyr	Arg	Thr	Phe	Phe 30	Ala	Leu
5	Lys	Gly	Leu 35	Thr	Thr	Ser	Arg	Gly 40	Glu	Pro	Val	Gln	Ala 45	Val	Tyr	Gly
LO	Phe	Ala 50	Lys	Ser	Leu	Leu	Lys 55	Ala	Leu	Lys	Glu	Asp 60	Gly	Tyr	Lys	Ala
	Val 65	Phe	Val	Val	Phe	Asp 70	Ala	Lys	Ala	Pro	Ser 75	Phe	Arg	His	Glu	Ala 80
1.5	Tyr	Glu	Ala	Tyr	Lys 85	Ala	Gly	Arg	Ala	Pro 90	Thr	Pro	Glu	Asp	Phe 95	Pro
2.0	Arg	Gln	Leu	Ala 100	Leu	Ile	Lys	Glu	Leu 105	Val	Asp	Leu	Leu	Gly 110	Phe	Thr
20	Arg	Leu	Glu 115	Val	Pro	Gly	Tyr	Glu 120	Ala	Asp	Asp	Val	Leu 125	Ala	Thr	Leu
25	Ala	Lys 130	Lys	Ala	Glu	Lys	Glu 135	Gly	Tyr	Glu	Val	Arg 140	Ile	Leu	Thr	Ala
	Asp 145	Arg	Asp	Leu	Tyr	Gln 150	Leu	Val	Ser	Asp	Arg 155	Val	Ala	Val	Leu	His
30	Pro	Glu	Gly	His	Leu 165	Ile	Thr	Pro	Glu	Trp 170	Leu	Trp	Glu	Lys	Tyr 175	Gly
2 E	Leu	Arg	Pro	Glu 180	Gln	Trp	Val	Asp	Phe 185	Arg	Ala	Leu	Val	Gly 190	Asp	Pro
	Ser	Asp	Asn 195	Leu	Pro	Gly	Val	Lys 200	Gly	Ile	Gly	Glu	Lys 205	Thr	Ala	Leu
40	Lys	Leu 210	Leu	Lys	Glu	Trp	Gly 215	Ser	Leu	Glu	Asn	Leu 220	Leu	Lys	Asn	Leu
	Asp 225	Arg	Val	Lys	Pro	Glu 230	Asn	Val	Arg	Glu	Lys 235	Ile	Lys	Ala	His	Leu 240
45	Glu	Asp	Leu	Arg	Leu 245	Ser	Leu	Glu	Leu	Ser 250	Arg	Val	Arg	Thr	Asp 255	Leu
50	Pro	Leu	Glu	Val 260	Asp	Leu	Ala	Gln	Gly 265	Arg	Glu	Pro	Asp	Arg 270	Glu	GlΣ
30	Leu	Arg	Ala 275		Leu	Glu	Arg	Leu 280	Glu	Phe	Gly	Ser	Leu 285	Leu	His	Glı
55	Phe	Gly 290	Leu	Leu	Glu	Ala	Pro 295		Pro	Leu	Glu	Glu 300	Ala	Pro	Trp	Pro
	Pro 305		Glu	Gly	Ala	Phe 310	Val	Gly	Phe	Val	Leu 315	Ser	Arg	Pro	Glu	Pro 320
60	Met	Trp	Ala	Glu	Leu 325		Ala	Leu	Ala	Ala 330		Arg	Gly	Gly	Arg 335	Va.

	His	Arg	Ala	Ala 340	Asp	Pro	Leu	Ala	Gly 345	Leu	Lys	Asp	Leu	Lys 350	Glu	Val
5	Arg	Gly	Leu 355	Leu	Ala	Lys	Asp	Leu 360	Ala	Val	Leu	Ala	Ser 365	Arg	Glu	Gly
	Leu	Asp 370	Leu	Val	Pro	Gly	Asp 375	Asp	Pro	Met	Leu	Leu 380	Ala	Tyr	Leu	Leu
10	Asp 385	Pro	Ser	Asn	Thr	Thr 390	Pro	Glu	Gly	Val	Ala 395	Arg	Arg	Tyr	Gly	Gly 400
<b>1</b> =	Glu	Trp	Thr	Glu	Asp 405	Ala	Ala	His	Arg	Ala 410	Leu	Leu	Ser	Glu	Arg 415	Leu
15	His	Arg	Asn	Leu 420	Leu	Lys	Arg	Leu	Glu 425	Gly	Glu	Glu	Lys	Leu 430	Leu	Trp
20	Leu	Tyr	His 435	Glu	Val	Glu	Lys	Pro 440	Leu	Ser	Arg	Val	Leu 445	Ala	His	Met
	Glu	Ala 450	Thr	Gly	Val	Arg	Arg 455	Asp	Val	Ala	Tyr	Leu 460	Gln	Ala	Leu	Ser
25	Leu 465	Glu	Leu	Ala	Glu	Glu 470	Ile	Arg	Arg	Leu	Glu 475	Glu	Glu	Val	Phe	Arg 480
diam than the	Leu	Ala	Gly	His	Pro 485	Phe	Asn	Leu	Asn	Ser 490	Arg	Asp	Gln	Leu	Glu 495	Arg
30	Val	Leu	Phe	Asp 500	Glu	Leu	Arg	Leu	Pro 505	Ala	Leu	Gly	Lys	Thr 510	Gln	Lys
35	Thr	Gly	Lys 515	Arg	Ser	Thr	Ser	Ala 520	Ala	Val	Leu	Glu	Ala 525	Leu	Arg	Glu
Marco Marco	Ala	His 530	Pro	Ile	Val	Glu	Lys 535	Ile	Leu	Gln	His	Arg 540	Glu	Leu	Thr	Lys
40	Leu 545	Lys	Asn	Thr	Tyr	Val 550	Asp	Pro	Leu	Pro	Ser 555	Leu	Val	His	Pro	Arg 560
i sā	Thr	Gly	Arg	Leu	His 565		Arg	Phe	Asn	Gln 570	Thr	Ala	Thr	Ala	Thr 575	Gly
45	Arg	Leu	Ser	Ser 580	Ser	Asp	Pro	Asn	Leu 585	Gln	Asn	Ile	Pro	Val 590	Arg	Thr
50	Pro	Leu	Gly 595	Gln	Arg	Ile	Arg	Arg 600	Ala	Phe	Val	Ala	Glu 605	Ala	Gly	Trp
	Ala	Leu 610	Val	Ala	Leu	Asp	Tyr 615	Ser	Gln	Ile	Glu	Leu 620	Arg	Val	Leu	Ala
55	His 625		Ser	Gly	Asp	Glu 630	Asn	Leu	Ile	Arg	Val 635	Phe	Gln	Glu	Gly	Lys 640
60	Asp	Ile	His	Thr	Gln 645	Thr	Ala	Ser	Trp	Met 650	Phe	Gly	Val	Pro	Pro 655	Glu
33	Ala	Val	Asp	Pro 660		Met	Arg	Arg	Ala 665		Lys	Thr	Val	Asn 670	Phe	Gly

	Val	Leu	Tyr 675	Gly	Met	Ser	Ala	His 680	Arg	Leu	Ser	Gln	Glu 685	Leu	Ala	Ile	
5	Pro	Tyr 690	Glu	Glu	Ala	Val	Ala 695	Phe	Ile	Glu	Arg	Tyr 700	Phe	Gln	Ser	Phe	
	Pro 705	Lys	Val	Arg	Ala	Trp 710	Ile	Glu	Lys	Thr	Leu 715	Glu	Glu	Gly	Arg	Lys 720	
10	Arg	Gly	Tyr	Val	Glu 725	Thr	Leu	Phe	Gly	Arg 730	Arg	Arg	Tyr	Val	Pro 735	Asp	
	Leu	Asn	Ala	Arg 740	Val	Lys	Ser	Val	Arg 745	Glu	Ala	Ala	Glu	Arg 750	Met	Ala	
15	Phe	Asn	Met 755	Pro	Val	Gln	Gly	Thr 760	Ala	Ala	Asp	Leu	Met 765	Lys	Leu	Ala	
20	Met	Val 770	Lys	Leu	Phe	Pro	Arg 775	Leu	Arg	Glu	Met	Gly 780	Ala	Arg	Met	Leu	
	Leu 785	Gln	Val	His	Asn	Glu 790	Leu	Leu	Leu	Glu	Ala 795	Pro	Gln	Ala	Arg	Ala 800	
25	Glu	Glu	Val	Ala	Ala 805	Leu	Ala	Lys	Glu	Ala 810	Met	Glu	Lys	Ala	Tyr 815	Pro	
	Leu	Ala	Val	Pro 820	Leu	Glu	Val	Glu	Val 825	Gly	Met	Gly	Glu	Asp 830	Trp	Leu	
25	Ser	Ala	Lys 835	Gly	His	His	His	His 840	His	His							
35	<210	0 >	21														
	<21	1>	30														
100 M	<21	2>	DNA														
4 O was the	<21	3>	synt	heti	С												
	<40		21														
45	cag	gagg	agc	tcgt	tggc	ga c	ctgg	agga	g								30
	<21	0>	22														
50	<21	1>	2526														
	<21	2>	DNA														
	<21	3>	Ther	mus	ther	moph	ilus										
55	<40	0>	22														
	atg	aatt	ccg	aggc	gatg	ct t	ccgc	tctt	t ga	accc	aaag	gcc	gggt	cct	cctg	gtggac	60
60	ggc	cacc	acc	tgga	ctac	cg c	acct	tctt	c go	cctg	aagg	gco	tcac	cac	gagc	cggggc	120
	gaa	.ccgg	ıtgc	aggo	ggtc	ta c	ggct	tcgc	c aa	.gagc	ctcc	tca	.aggc	cct	gaag	gaggac	180

gggtacaagg	ccgtcttcgt	ggtctttgac	gccaaggccc	cctccttccg	ccacgaggcc	240
tacgaggcct	acaaggcggg	gagggccccg	acccccgagg	acttcccccg	gcagctcgcc	300
ctcatcaagg	agctggtgga	cctcctgggg	tttacccgcc	tcgaggtccc	cggctacgag	360
gcggacgacg	ttctcgccac	cctggccaag	aaggcggaaa	aggaggggta	cgaggtgcgc	420
atcctcaccg	ccgaccgcga	cctctaccaa	ctcgtctccg	accgcgtcgc	cgtcctccac	480
cccgagggcc	acctcatcac	cccggagtgg	ctttgggaga	agtacggcct	caggccggag	540
cagtgggtgg	acttccgcgc	cctcgtgggg	gacccctccg	acaacctccc	cggggtcaag	600
ggcatcgggg	agaagaccgc	cctcaagctc	ctcaaggagt	ggggaagcct	ggaaaacctc	660
ctcaagaacc	tggaccgggt	aaagccagaa	aacgtccggg	agaagatcaa	ggcccacctg	720
gaagacctca	ggctctcctt	ggagctctcc	cgggtgcgca	ccgacctccc	cctggaggtg	780
gacctcgccc	aggggcggga	gcccgaccgg	gaggggctta	gggccttcct	ggagaggctg	840
gagttcggca	gcctcctcca	cgagttcggc	ctcctggagg	ccccgcccc	cctggaggag	900
gccccctggc	ccccgccgga	aggggccttc	gtgggcttcg	tecteteceg	ccccgagccc	960
atgtgggcgg	agcttaaagc	cctggccgcc	tgcaggggcg	gccgcgtgca	ccgggcagca	1020
gaccccttgg	cggggctaaa	ggacctcaag	gaggtccggg	gcctcctcgc	caaggacctc	1080
gccgtcttgg	cctcgaggga	ggggctagac	ctcgtgcccg	gggacgaccc	catgctcctc	1140
gcctacctcc	tggacccttc	gaacaccacc	cccgaggggg	tggcgcggcg	ctacgggggg	1200
gagtggacgg	aggacgccgc	ccaccgggcc	ctcctctcgg	agaggctcca	tcggaacctc	1260
cttaagcgcc	tcgaggggga	ggagaagctc	ctttggctct	accacgaggt	ggaaaagccc	1320
ctctcccggg	tcctggccca	tatggaggcc	accggggtac	ggcgggacgt	ggcctacctt	1380
caggcccttt	ccctggagct	tgcggaggag	atccgccgcc	tcgaggagga	ggtcttccgc	1440
ttggcgggcc	accccttcaa	cctcaactcc	cgggaccagc	tggaaagggt	gctctttgac	1500
gagcttaggc	ttcccgcctt	ggggaagacg	caaaagacag	gcaagcgctc	caccagcgcc	1560
gcggtgctgg	aggccctacg	ggaggcccac	cccatcgtgg	agaagatcct	ccagcaccgg	1620
gagctcacca	agctcaagaa	cacctacgtg	gaccccctcc	caagcctcgt	ccacccgagg	1680
acgggccgcc	tccacacccg	cttcaaccag	acggccacgg	ccacggggag	gcttagtagc	1740
tccgacccca	acctgcagaa	catccccgtc	cgcaccccct	tgggccagag	gatccgccgg	1800
gccttcgtgg	ccgaggcggg	ttgggcgttg	gtggccctgg	actatagcca	gatagagctc	1860
cgcgtcctcg	cccacctctc	cggggacgaa	aacctgatca	gggtcttcca	ggaggggaag	1920
gacatccaca	cccagaccgc	aagctggatg	ttcggcgtcc	ccccggaggc	cgtggacccc	1980
ctgatgcgcc	gggcggccaa	gacggtgaac	ttcggcgtcc	tctacggcat	gtccgcccat	2040

	aggetetece aggagettge catecectae gaggaggegg tggeetttat agagegetae	2100
	ttccaaaget teeccaaggt gegggeetgg atagaaaaga eeetggagga ggggaggaag	2160
5	cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg	2220
	gtgaagageg teagggagge egeggagege atggeettea acatgeeegt eeagggeace	2280
	geegeegaee teatgaaget egeeatggtg aagetettee eeegeeteeg ggagatgggg	2340
10	gecegeatge tectecaggt egecaaegag etecteetgg aggeceecea agegegggee	2400
	gaggaggtgg eggetttgge caaggaggee atggagaagg eetateeeet egeegtgeee	2460
15	ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac	2520
	caccac	2526
20	<210> 23	
	<211> 842	
25	<212> PRT	
	<213> Thermus thermophilus	
	<400> 23	
30	Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val 1 5 10 15	
2 3 3 3 5 mg mg mg 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu 20 25 30	
40	Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly 35 40 45	
40	Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala 50 55 60	
	Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala 65 70 75 80	
45	Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro 85 90 95	
F.0	Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr 100 105 110	
50	Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu 115 120 125	
55	Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala 130 135 140	
	Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His 145 150 150	
60	Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly 165 170 175	

	Leu	Arg	Pro	Glu 180	Gln	Trp	Val	Asp	Phe 185	Arg	Ala	Leu	Val	Gly 190	Asp	Pro
5	Ser	Asp	Asn 195	Leu	Pro	Gly	Val	Lys 200	Gly	Ile	Gly	Glu	Lys 205	Thr	Ala	Leu
	Lys	Leu 210	Leu	Lys	Glu	Trp	Gly 215	Ser	Leu	Glu	Asn	Leu 220	Leu	Lys	Asn	Leu
10	Asp 225	Arg	Val	Lys	Pro	Glu 230	Asn	Val	Arg	Glu	Lys 235	Ile	Lys	Ala	His	Leu 240
1 5	Glu	Asp	Leu	Arg	Leu 245	Ser	Leu	Glu	Leu	Ser 250	Arg	Val	Arg	Thr	Asp 255	Leu
15	Pro	Leu	Glu	Val 260	Asp	Leu	Ala	Gln	Gly 265	Arg	Glu	Pro	Asp	Arg 270	Glu	Gly
20	Leu	Arg	Ala 275	Phe	Leu	Glu	Arg	Leu 280	Glu	Phe	Gly	Ser	Leu 285	Leu	His	Glu
	Phe	Gly 290	Leu	Leu	Glu	Ala	Pro 295	Ala	Pro	Leu	Glu	Glu 300	Ala	Pro	Trp	Pro
25	Pro 305	Pro	Glu	Gly	Ala	Phe 310	Val	Gly	Phe	Val	Leu 315	Ser	Arg	Pro	Glu	Pro 320
	Met	Trp	Ala	Glu	Leu 325	Lys	Ala	Leu	Ala	Ala 330	Cys	Arg	Gly	Gly	Arg 335	Val
30.	His	Arg	Ala	Ala 340	Asp	Pro	Leu	Ala	Gly 345	Leu	Lys	Asp	Leu	Lys 350	Glu	Val
35	Arg	Gly	Leu 355	Leu	Ala	Lys	Asp	Leu 360	Ala	Val	Leu	Ala	Ser 365	Arg	Glu	Gly
		370					375					380			Leu	
40	Asp 385	Pro	Ser	Asn	Thr	Thr 390	Pro	Glu	Gly	Val	Ala 395	Arg	Arg	Tyr	Gly	Gly 400
45		_			405					410					Arg 415	
	His	Arg	Asn	Leu 420	Leu	Lys	Arg	Leu	Glu 425	Gly	Glu	Glu	Lys	Leu 430	Leu	Trp
50			435					440					445		His	
		450					455					460			Leu	
55	465					470					475					Arg 480
60					485					490					Glu 495	
	Val	Leu	Phe	Asp 500	Glu	Leu	Arg	Leu	Pro 505	Ala	Leu	Gly	Lys	Thr 510	Gln	Lys

	Thr	Gly	Lys 515	Arg	Ser	Thr	Ser	Ala 520	Ala	Val	Leu	Glu	Ala 525	Leu	Arg	Glu
5	Ala	His 530	Pro	Ile	Val	Glu	Lys 535	Ile	Leu	Gln	His	Arg 540	Glu	Leu	Thr	Lys
	Leu 545	Lys	Asn	Thr	Tyr	Val 550	Asp	Pro	Leu	Pro	Ser 555	Leu	Val	His	Pro	Arg 560
10	Thr	Gly	Arg	Leu	His 565	Thr	Arg	Phe	Asn	Gln 570	Thr	Ala	Thr	Ala	Thr 575	Gly
15	Arg	Leu	Ser	Ser 580	Ser	Asp	Pro	Asn	Leu 585	Gln	Asn	Ile	Pro	Val 590	Arg	Thr
15	Pro	Leu	Gly 595	Gln	Arg	Ile	Arg	Arg 600	Ala	Phe	Val	Ala	Glu 605	Ala	Gly	Trp
20	Ala	Leu 610	Val	Ala	Leu	Asp	Tyr 615	Ser	Gln	Ile	Glu	Leu 620	Arg	Val	Leu	Ala
	His 625	Leu	Ser	Gly	Asp	Glu 630	Asn	Leu	Ile	Arg	Val 635	Phe	Gln	Glu	Gly	Lys 640
25	Asp	Ile	His	Thr	Gln 645	Thr	Ala	Ser	Trp	Met 650	Phe	Gly	Val	Pro	Pro 655	Glu
30	Ala	Val	Asp	Pro 660	Leu	Met	Arg	Arg	Ala 665	Ala	Lys	Thr	Val	Asn 670	Phe	Gly
30	Val	Leu	Tyr 675	Gly	Met	Ser	Ala	His 680	Arg	Leu	Ser	Gln	Glu 685	Leu	Ala	Ile
35	Pro	Tyr 690	Glu	Glu	Ala	Val	Ala 695	Phe	Ile	Glu	Arg	Tyr 700	Phe	Gln	Ser	Phe
	Pro 705	Lys	Val	Arg	Ala	Trp 710	Ile	Glu	Lys	Thr	Leu 715	Glu	Glu	Gly	Arg	Lys 720
40	Arg	Gly	Tyr	Val	Glu 725	Thr	Leu	Phe	Gly	Arg 730	Arg	Arg	Tyr	Val	Pro 735	Asp
45	Leu	Asn	Ala	Arg 740	Val	Lys	Ser		Arg 745		Ala	Ala	Glu	Arg 750	Met	Ala
±3	Phe	Asn	Met 755	Pro	Val	Gln	Gly	Thr 760	Ala	Ala	Asp	Leu	Met 765	Lys	Leu	Ala
50	Met	Val 770	Lys	Leu	Phe	Pro	Arg 775	Leu	Arg	Glu	Met	Gly 780	Ala	Arg	Met	Leu
	Leu 785	Gln	Val	Ala	Asn	Glu 790	Leu	Leu	Leu	Glu	Ala 795	Pro	Gln	Ala	Arg	Ala 800
55	Glu	Glu	Val	Ala	Ala 805	Leu	Ala	Lys	Glu	Ala 810		Glu	Lys	Ala	Tyr 815	Pro
60	Leu	Ala	Val	Pro 820	Leu	Glu	Val	Glu	Val 825	Gly	Met	Gly	Glu	Asp 830	Trp	Leu
•	Ser	Ala	Lys 835		His	His	His	His 840		His						

```
<210>
             24
             30
      <211>
5
      <212>
             DNA
      <213>
             synthetic
      <400>
10
                                                                               30
      caggaggagc tcgttggcga cctggaggag
15
       <210> 25
             2526
       <211>
       <212>
             DNA
20
             Thermus thermophilus
             25
       <400>
                                                                               60
       atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac
25 3 3 3 3 4
                                                                              120
       ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc
       gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac
                                                                              180
       gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttccg ccacgaggcc
                                                                              240
       tacgaggeet acaaggeggg gagggeeeeg acceeegagg actteeeeeg geagetegee
                                                                              300
       ctcatcaagg agctggtgga cctcctgggg tttacccgcc tcgaggtccc cggctacgag
                                                                              360
       gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc
                                                                              420
                                                                              480
       atceteaceg cegacegega cetetaceaa etegteteeg acegegtege egtecteeac
       cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag
                                                                              540
       cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag
                                                                              600
 1.4
                                                                              660
       ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc
45
                                                                              720
       ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg
       gaagacetea ggeteteett ggagetetee egggtgegea eegaeeteee eetggaggtg
                                                                              780
50
       gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg
                                                                              840
                                                                              900
       gagttcggca gcctcctcca cgagttcggc ctcctggagg cccccgcccc cctggaggag
       gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc
                                                                              960
55
                                                                             1020
       atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca
                                                                             1080
       gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc
60
                                                                             1140
       gccgtcttgg cctcgaggga ggggctagac ctcgtgcccg gggacgaccc catgctcctc
```

```
1200
      gectacetee tggaceette gaacaceace eeegaggggg tggegeggeg etacgggggg
                                                                            1260
      gagtggacgg aggacgccgc ccaccgggcc ctcctctcgg agaggctcca tcggaacctc
                                                                            1320
      cttaagcgcc tcgagggga ggagaagctc ctttggctct accacgaggt ggaaaagccc
5
                                                                            1380
      ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt
      caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc
                                                                            1440
10
      ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac
                                                                            1500
      gagettagge ttecegeett gaagaagaeg aagaagaeag geaagegete caceagegee
                                                                            1560
                                                                            1620
      geggtgetgg aggecetacg ggaggeceae eccategtgg agaagateet eeageaeegg
15
                                                                            1680
      gageteacea ageteaagaa cacetaegtg gaceceetee caageetegt ecaceegagg
                                                                            1740
       acgggccgcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc
20
                                                                            1800
       tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg
       gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc
                                                                            1860
                                                                            1920
       cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggtcttcca ggaggggaag
1980
       gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccggaggc cgtggacccc
       ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccat
                                                                            2040
                                                                            2100
       aggetetece aggagettge catecectae gaggaggegg tggeetttat agagegetae
                                                                            2160
       ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag
       cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg
                                                                            2220
                                                                            2280
       gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc
       geegeegace teatgaaget egecatggtg aagetettee eeegeeteeg ggagatgggg
                                                                            2340
       gecegeatge tectecaggt egecaaegag etecteetgg aggeeecea agegegggee
                                                                            2400
                                                                            2460
       gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac
                                                                            2520
45
                                                                            2526
       caccac
              26
50
       <210>
       <211>
              842
       <212>
              PRT
55
              Thermus thermophilus
       <213>
       <400>
       Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
60
```

	Leu	Leu	Val	Asp 20	Gly	His	His	Leu	Ala 25	Tyr	Arg	Thr	Phe	Phe 30	Ala	Leu
5	Lys	Gly	Leu 35	Thr	Thr	Ser	Arg	Gly 40	Glu	Pro	Val	Gln	Ala 45	Val	Tyr	Gly
	Phe	Ala 50	Lys	Ser	Leu	Leu	Lys 55	Ala	Leu	Lys	Glu	Asp 60	Gly	Tyr	Lys	Ala
10	Val 65	Phe	Val	Val	Phe	Asp 70	Ala	Lys	Ala	Pro	Ser 75	Phe	Arg	His	Glu	Ala 80
1 5	Tyr	Glu	Ala	Tyr	Lys 85	Ala	Gly	Arg	Ala	Pro 90	Thr	Pro	Glu	Asp	Phe 95	Pro
15	Arg	Gln	Leu	Ala 100	Leu	Ile	Lys	Glu	Leu 105	Val	Asp	Leu	Leu	Gly 110	Phe	Thr
20	Arg	Leu	Glu 115	Val	Pro	Gly	Tyr	Glu 120	Ala	Asp	Asp	Val	Leu 125	Ala	Thr	Leu
	Ala	Lys 130	Lys	Ala	Glu	Lys	Glu 135	Gly	Tyr	Glu	Val	Arg 140	Ile	Leu	Thr	Ala
25	Asp 145	Arg	Asp	Leu	Tyr	Gln 150	Leu	Val	Ser	Asp	Arg 155	Val	Ala	Val	Leu	His 160
Bern Min Will	Pro	Glu	Gly	His	Leu 165	Ile	Thr	Pro	Glu	Trp 170	Leu	Trp	Glu	Lys	Tyr 175	Gly
30	Leu	Arg	Pro	Glu 180	Gln	Trp	Val	Asp	Phe 185	Arg	Ala	Leu	Val	Gly 190	Asp	Pro
35	Ser	Asp	Asn 195	Leu	Pro	Gly	Val	Lys 200	Gly	Ile	Gly	Glu	Lys 205	Thr	Ala	Leu
Manual States	Lys	Leu 210	Leu	Lys	Glu	Trp	Gly 215	Ser	Leu	Glu	Asn	Leu 220	Leu	Lys	Asn	Leu
40	Asp 225	Arg	Val	Lys	Pro	Glu 230	Asn	Val	Arg	Glu	Lys 235	Ile	Lys	Ala	His	Leu 240
45	Glu	Asp	Leu	Arg	Leu 245	Ser	Leu	Glu	Leu	Ser 250	Arg	Val	Arg	Thr	Asp 255	Leu
10	Pro	Leu	Glu	Val 260	Asp	Leu	Ala	Gln	Gly 265	Arg	Glu	Pro	Asp	Arg 270	Glu	Gly
50			275					280					285			Glu
	Phe	Gly 290	Leu	Leu	Glu	Ala	Pro 295	Ala	Pro	Leu	Glu	Glu 300	Ala	Pro	Trp	Pro
55	305					310					315					Pro 320
60					325					330					Arg 335	
	His	Arg	Ala	Ala 340	Asp	Pro	Leu	Ala	Gly 345		Lys	Asp	Leu	Lys 350	Glu	Val

	Arg	Gly	Leu 355	Leu	Ala	Lys	Asp	Leu 360	Ala	Val	Leu	Ala	Ser 365	Arg	Glu	Gly
5	Leu	Asp 370	Leu	Val	Pro	Gly	Asp 375	Asp	Pro	Met	Leu	Leu 380	Ala	Tyr	Leu	Leu
	Asp 385	Pro	Ser	Asn	Thr	Thr 390	Pro	Glu	Gly	Val	Ala 395	Arg	Arg	Tyr	Gly	Gly 400
10	Glu	Trp	Thr	Glu	Asp 405	Ala	Ala	His	Arg	Ala 410	Leu	Leu	Ser	Glu	Arg 415	Leu
a =	His	Arg	Asn	Leu 420	Leu	Lys	Arg	Leu	Glu 425	Gly	Glu	Glu	Lys	Leu 430	Leu	Trp
15	Leu	Tyr	His 435	Glu	Val	Glu	Lys	Pro 440	Leu	Ser	Arg	Val	Leu 445	Ala	His	Met
20	Glu	Ala 450	Thr	Gly	Val	Arg	Arg 455	Asp	Val	Ala	Tyr	Leu 460	Gln	Ala	Leu	Ser
	Leu 465	Glu	Leu	Ala	Glu	Glu 470	Ile	Arg	Arg	Leu	Glu 475	Glu	Glu	Val	Phe	Arg 480
25	Leu	Ala	Gly	His	Pro 485	Phe	Asn	Leu	Asn	Ser 490	Arg	Asp	Gln	Leu	Glu 495	Arg
dian than And	Val	Leu	Phe	Asp 500	Glu	Leu	Arg	Leu	Pro 505	Ala	Leu	Lys	Lys	Thr 510	Lys	Lys
30	Thr	Gly	Lys 515	Arg	Ser	Thr	Ser	Ala 520	Ala	Val	Leu	Glu	Ala 525	Leu	Arg	Glu
35	Ala	His 530	Pro	Ile	Val	Glu	Lys 535	Ile	Leu	Gln	His	Arg 540	Glu	Leu	Thr	Lys
Marie anna Marie Marie Marie anna	Leu 545	Lys	Asn	Thr	Tyr	Val 550	Asp	Pro	Leu	Pro	Ser 555	Leu	Val	His	Pro	Arg 560
40	Thr	Gly	Arg	Leu	His 565	Thr	Arg	Phe	Asn	Gln 570	Thr	Ala	Thr	Ala	Thr 575	Gly
45	Arg	Leu	Ser	Ser 580		Asp	Pro	Asn	Leu 585	Gln	Asn	Ile	Pro	Val 590	Arg	Thr
43	Pro	Leu	Gly 595	Gln	Arg	Ile	Arg	Arg 600	Ala	Phe	Val	Ala	Glu 605	Ala	Gly	Trp
50	Ala	Leu 610	Val	Ala	Leu	Asp	Tyr 615	Ser	Gln	Ile	Glu	Leu 620	Arg	Val	Leu	Ala
	His 625	Leu	Ser	Gly	Asp	Glu 630	Asn	Leu	Ile	Arg	Val 635	Phe	Gln	Glu	Gly	Lys 640
55	Asp	Ile	His	Thr	Gln 645	Thr	Ala	Ser	Trp	Met 650	Phe	Gly	Val	Pro	Pro 655	Glu
60	Ala	Val	Asp	Pro 660		Met	Arg	Arg	Ala 665		Lys	Thr	Val	Asn 670	Phe	Gly
	Val	Leu	Tyr 675		Met	Ser	Ala	His 680	Arg	Leu	Ser	Gln	Glu 685	. Leu	Ala	Ile

	Pro	Tyr 690	Glu	Glu	Ala	Val	Ala 695	Phe	Ile	Glu	Arg	Tyr 700	Phe	Gln	Ser	Phe
5	Pro 705	Lys	Val	Arg	Ala	Trp 710	Ile	Glu	Lys	Thr	Leu 715	Glu	Glu	Gly	Arg	Lys 720
	Arg	Gly	Tyr	Val	Glu 725	Thr	Leu	Phe	Gly	Arg 730	Arg	Arg	Tyr	Val	Pro 735	Asp
10	Leu	Asn	Ala	Arg 740	Val	Lys	Ser	Val	Arg 745	Glu	Ala	Ala	Glu	Arg 750	Met	Ala
15	Phe	Asn	Met 755	Pro	Val	Gln	Gly	Thr 760	Ala	Ala	Asp	Leu	Met 765	Lys	Leu	Ala
13	Met	Val 770	Lys	Leu	Phe	Pro	Arg 775	Leu	Arg	Glu	Met	Gly 780	Ala	Arg	Met	Leu
20	Leu 785	Gln	Val	Ala	Asn	Glu 790	Leu	Leu	Leu	Glu	Ala 795	Pro	Gln	Ala	Arg	Ala 800
	Glu	Glu	Val	Ala	Ala 805	Leu	Ala	Lys	Glu	Ala 810	Met	Glu	Lys	Ala	Tyr 815	Pro
25	Leu	Ala	Val	Pro 820	Leu	Glu	Val	Glu	Val 825	Gly	Met	Gly	Glu	Asp 830	Trp	Leu
3 0, 10, 10, 10, 10, 10, 10, 10, 10, 10,	Ser	Ala	Lys 835	Gly	His	His	His	His 840	His	His						
	<210	0 > :	27													
35	<21	1>	340													
21	<212	2> :	PRT													
The state of the s	<21	3>	Pyro	cocci	us fi	urio	sus									
40	<40	0>	27													
	Met 1	Gly	Val	Pro	Ile 5	Gly	Glu	Ile	Ile	Pro 10	Arg	Lys	Glu	Ile	Glu 15	Leu
45	Glu	Asn	Leu	Tyr 20	Gly	Lys	Lys	Ile	Ala 25	Ile	Asp	Ala	Leu	Asn 30	Ala	Ile
50	Tyr	Gln	Phe 35	Leu	Ser	Thr	Ile	Arg 40	Gln	Lys	Asp	Gly	Thr 45	Pro	Leu	Met
50	Asp	Ser 50	Lys	Gly	Arg	Ile	Thr 55	Ser	His	Leu	Ser	Gly 60	Leu	Phe	Tyr	Arg
55	Thr 65	Ile	Asn	Leu	Met	Glu 70		Gly	Ile	Lys	Pro 75	Val	Tyr	Val	Phe	qaA 08
	Gly	Glu	Pro	Pro	Glu 85	Phe	Lys	Lys	Lys	Glu 90	Leu	Glu	Lys	Arg	Arg 95	Glu
60	Ala	Arg	Glu	Glu 100		Glu	Glu	Lys	Trp 105		Glu	Ala	Leu	Glu 110	Lys	Gly

	Glu	Ile	Glu 1 <b>1</b> 5	Glu	Ala	Arg	Lys	Tyr 120	Ala	Gln	Arg	Ala	Thr 125	Arg	Val	Asn
, 5	Glu	Met 130	Leu	Ile	Glu	Asp	Ala 135	Lys	Lys	Leu	Leu	Glu 140	Leu	Met	Gly	Ile
	Pro 145	Ile	Val	Gln	Ala	Pro 150	Ser	Glu	Gly	Glu	Ala 155	Gln	Ala	Ala	Tyr	Met 160
10	Ala	Ala	Lys	Gly	Ser 165	Val	Tyr	Ala	Ser	Ala 170	Ser	Gln	Asp	Tyr	Asp 175	Ser
	Leu	Leu	Phe	Gly 180	Ala	Pro	Arg	Leu	Val 185	Arg	Asn	Leu	Thr	Ile 190	Thr	Gly
15	Lys	Arg	Lys 195	Leu	Pro	Gly	Lys	Asn 200	Val	Tyr	Val	Glu	Ile 205	Lys	Pro	Glu
20	Leu	Ile 210	Ile	Leu	Glu	Glu	Val 215	Leu	Lys	Glu	Leu	Lys 220	Leu	Thr	Arg	Glu
	Lys 225	Leu	Ile	Glu	Leu	Ala 230	Ile	Leu	Val	Gly	Thr 235	Asp	Tyr	Asn	Pro	Gly 240
100	Gly	Ile	Lys	Gly	Ile 245	Gly	Leu	Lys	Lys	Ala 250	Leu	Glu	Ile	Val	Arg 255	His
Seat Steen Steen	Ser	Lys	Asp	Pro 260	Leu	Ala	Lys	Phe	Gln 265	Lys	Gln	Ser	Asp	Val 270	Asp	Leu
3 O	Tyr	Ala	Ile 275	Lys	Glu	Phe	Phe	Leu 280	Asn	Pro	Pro	Val	Thr 285	Asp	Asn	Tyr
35	Asn	Leu 290	Val	Trp	Arg	Asp	Pro 295	Asp	Glu	Glu	Gly	Ile 300	Leu	Lys	Phe	Leu
Money Section	Cys 305	Asp	Glu	His	Asp	Phe 310	Ser	Glu	Glu	Arg	Val 315	Lys	Asn	Gly	Leu	Glu 320
40	Arg	Leu	Lys	Lys	Ala 325	Ile	Lys	Ser	Gly	Lys 330	Gln	Ser	Thr	Leu	Glu 335	Ser
	Trp	Phe	Lys	Arg 340												
45	<210	0 > 2	28	, 10												
	<21	1> :	326													
F0	<212	2> :	PRT													
50	<213	3 > 1	Metha	anoc	occu	s ja:	nnas	chii								
	<40	0>	28													
55	Met 1	Gly	Val	Gln	Phe 5	Gly	Asp	Phe	Ile	Pro 10	Lys	Asn	Ile	Ile	Ser 15	Phe
<b>50</b>	Glu	Asp	Leu	Lys 20	Gly	Lys	Lys	Val	Ala 25	Ile	Asp	Gly	Met	Asn 30	Ala	Leu
60	Tyr	Gln	Phe 35	Leu	Thr	Ser	Ile	Arg 40	Leu	Arg	Asp	Gly	Ser 45	Pro	Leu	Arg

	Asn	Arg 50	Lys	Gly	Glu	Ile	Thr 55	Ser	Ala	Tyr	Asn	Gly 60	Val	Phe	Tyr	Lys
5	Thr 65	Ile	His	Leu	Leu	Glu 70	Asn	Asp	Ile	Thr	Pro 75	Ile	Trp	Val	Phe	Asp 80
	Gly	Glu	Pro	Pro	Lys 85	Leu	Lys	Glu	Lys	Thr 90	Arg	Lys	Val	Arg	Arg 95	Glu
10	Met	Lys	Glu	Lys 100	Ala	Glu	Leu	Lys	Met 105	Lys	Glu	Ala	Ile	Lys 110	Lys	Glu
15	Asp	Phe	Glu 115	Glu	Ala	Ala	Lys	Tyr 120	Ala	Lys	Arg	Val	Ser 125	Tyr	Leu	Thr
	Pro	Lys 130	Met	Val	Glu	Asn	Cys 135	Lys	Tyr	Leu	Leu	Ser 140	Leu	Met	Gly	Ile
20	Pro 145	Tyr	Val	Glu	Ala	Pro 150	Ser	Glu	Gly	Glu	Ala 155	Gln	Ala	Ser	Tyr	Met 160
	Ala	Lys	Lys	Gly	Asp 165	Val	Trp	Ala	Val	Val 170	Ser	Gln	Asp	Tyr	Asp 175	Ala
25	Leu	Leu	Tyr	Gly 180	Ala	Pro	Arg	Val	Val 185	Arg	Asn	Leu	Thr	Thr 190	Thr	Lys
	Glu	Met	Pro 195	Glu	Leu	Ile	Glu	Leu 200	Asn	Glu	Val	Leu	Glu 205	Asp	Leu	Arg
	Ile	Ser 210	Leu	Asp	Asp	Leu	Ile 215	Asp	Ile	Ala	Ile	Phe 220	Met	Gly	Thr	Asp
	Tyr 225	Asn	Pro	Gly	Gly	Val 230	Lys	Gly	Ile	Gly	Phe 235	Lys	Arg	Ala	Tyr	Glu 240
	Leu	Val	Arg	Ser	Gly 245	Val	Ala	Lys	Asp	Val 250	Leu	Lys	Lys	Glu	Val 255	Glu
40	Tyr	Tyr	Asp	Glu 260	Ile	Lys	Arg	Ile	Phe 265	Lys	Glu	Pro	Lys	Val 270	Thr	Asp
45	Asn	Tyr	Ser 275	Leu	Ser	Leu	Lys	Leu 280	Pro	Asp	Lys	Glu	Gly 285	Ile	Ile	Lys
	Phe	Leu 290	Val	Asp	Glu	Asn	Asp 295	Phe	Asn	Tyr	Asp	Arg 300	Val	Lys	Lys	His
50	Val 305	_	Lys	Leu	Tyr	Asn 310	Leu	Ile	Ala	Asn	Lys 315	Thr	Lys	Gln	Lys	Thr 320
	Leu	Asp	Ala	Trp	Phe 325	Lys										
55	<21	0>	29													
	<21	<211> 328														
60	<21	2>	PRT													
	<21	3>	Meth	anob	acte	rium	the	rmoa	utot	roph	icum					

Met Gly Val Lys Leu Arg Asp Val Val Ser Pro Arg Arg Ile Arg Leu 5 Glu Asp Leu Arg Gly Arg Thr Val Ala Val Asp Ala Ala Asn Thr Leu Tyr Gln Phe Leu Ser Ser Ile Arg Gln Arg Asp Gly Thr Pro Leu Met 10 Asp Ser Arg Gly Arg Val Thr Ser His Leu Ser Gly Ile Leu Tyr Arg Thr Ala Ala Val Met Glu Arg Glu Ile Arg Val Ile Tyr Val Phe Asp 15 Gly Arg Ser His His Leu Lys Gly Glu Thr Val Ser Arg Arg Ala Asp 20 Ile Arg Lys Lys Ser Glu Val Glu Trp Lys Arg Ala Leu Glu Glu Gly Asp Ile Asp Arg Ala Arg Lys Tyr Ala Val Arg Ser Ser Arg Met Ser 30 35 Ser Glu Ile Leu Glu Ser Ser Lys Arg Leu Leu Glu Leu Leu Gly Ile Pro Tyr Val Gln Ala Pro Gly Glu Gly Glu Ala Gln Ala Ser Tyr Met Val Lys Met Gly Asp Ala Trp Ala Val Ala Ser Gln Asp Tyr Asp Cys Leu Leu Phe Gly Ala Pro Arg Val Val Arg Lys Val Thr Leu Ser Gly 40 185 Lys Leu Glu Asp Pro His Ile Ile Glu Leu Glu Ser Thr Leu Arg Ala Leu Ser Ile Ser His Thr Gln Leu Val Asp Met Ala Leu Leu Val Gly 215 i så Thr Asp Phe Asn Glu Gly Val Lys Gly Tyr Gly Ala Arg Arg Gly Leu 45 Lys Leu Ile Arg Glu Lys Gly Asp Ile Phe Lys Val Ile Arg Asp Leu 50 Glu Ala Asp Ile Gly Gly Asp Pro Gln Val Leu Arg Arg Ile Phe Leu Glu Pro Glu Val Ser Glu Asp Tyr Glu Ile Arg Trp Arg Lys Pro Asp 55 Val Glu Gly Val Ile Glu Phe Leu Cys Thr Glu His Gly Phe Ser Glu Asp Arg Val Arg Asp Ala Leu Lys Lys Phe Glu Gly Ala Ser Ser Thr 60 315

```
5
       <210> 30
       <211>
             336
10
       <212>
             PRT
       <213>
            Afu
       <400> 30
15
       Met Gly Ala Asp Ile Gly Asp Leu Phe Glu Arg Glu Glu Val Glu Leu
       Glu Tyr Phe Ser Gly Lys Lys Ile Ala Val Asp Ala Phe Asn Thr Leu
20
       Tyr Gln Phe Ile Ser Ile Ile Arg Gln Pro Asp Gly Thr Pro Leu Lys
2 3 3 3 4
       Asp Ser Gln Gly Arg Ile Thr Ser His Leu Ser Gly Ile Leu Tyr Arg
                                55
       Val Ser Asn Met Val Glu Val Gly Ile Arg Pro Val Phe Val Phe Asp
       Gly Glu Pro Pro Glu Phe Lys Lys Ala Glu Ile Glu Glu Arg Lys Lys
       Arg Arg Ala Glu Ala Glu Glu Met Trp Ile Ala Ala Leu Gln Ala Gly
                                       105
       Asp Lys Asp Ala Lys Lys Tyr Ala Gln Ala Ala Gly Arg Val Asp Glu
       Tyr Ile Val Asp Ser Ala Lys Thr Leu Leu Ser Tyr Met Gly Ile Pro
×£
       Phe Val Asp Ala Pro Ser Glu Gly Glu Ala Gln Ala Ala Tyr Met Ala
45
                                                155
       Ala Lys Gly Asp Val Glu Tyr Thr Gly Ser Gln Asp Tyr Asp Ser Leu
       Leu Phe Gly Ser Pro Arg Leu Ala Arg Asn Leu Ala Ile Thr Gly Lys
50
       Arg Lys Leu Pro Gly Lys Asn Val Tyr Val Asp Val Lys Pro Glu Ile
55
       Ile Ile Leu Glu Ser Asn Leu Lys Arg Leu Gly Leu Thr Arg Glu Gln
       Leu Ile Asp Ile Ala Ile Leu Val Gly Thr Asp Tyr Asn Glu Gly Val
60
       Lys Gly Val Gly Val Lys Lys Ala Leu Asn Tyr Ile Lys Thr Tyr Gly
```

				245					250					255			
	Asp Ile	e Phe	Arg 260	Ala	Leu	Lys	Ala	Leu 265	Lys	Val	Asn	Ile	Asp 270	His	Val		
5	Glu Glu	1 Ile 275		Asn	Phe	Phe	Leu 280	Asn	Pro	Pro	Val	Thr 285	Asp	Asp	Tyr		
10	Arg Ile		Phe	Arg	Glu	Pro 295	Asp	Phe	Glu	Lys	Ala 300	Ile	Glu	Phe	Leu		
	Cys Glu 305	ı Glu	His	Asp	Phe 310	Ser	Arg	Glu	Arg	Val 315	Glu	Lys	Ala	Leu	Glu 320		
15	Lys Let	ı Lys	Ala	Leu 325	Lys	Ser	Thr	Gln	Ala 330	Thr	Leu	Glu	Arg	Trp 335	Phe		
20	<210>	31															
20	<211>	27															
	<212>	DNA															
25,	<213>	synt	heti	С													
ud U	<220><221> misc_feature<222> (17)(18)<223> n = degeneracy																
20	<220><221> misc_feature<222> (27)(28)<223> n = degeneracy																
	<400>	31															
	atctctagca ctgctgtntt ygayggn														27		
35																	
#:	<210> 32																
	<211> 31																
40	<212>	DNA															
	<213>	synt	heti	С													
<b>.</b> _	<220><	221>	mis	c_fe	atur	e<22	2>	(21)	(2	2)<2	23>	n =	deg	ener	асу		
45	<220><	221>	mis	c_fe	atur	e<22	2>	(27)	(2	8)<2	23>	n =	deg	ener	acy		
	<400>	32															
50	gatctc	tagc	actg	ctga	rg g	ngar	gcnc	a r									31
	<210>	33															
55	<211>	28															
	<212> DNA																
60	<213>	synt	heti	С													

```
28
      gatctctagc actgctcarg aytaygay
      <210> 34
5
      <211> 31
      <212> DNA
      <213> synthetic
10
      <220><221> misc_feature<222> (19)..(20)<223> n = degeneracy
      <220><221> misc_feature<222> (25)..(26)<223> n = degeneracy
15
      <400> 34
                                                                           31
      cttaaggtag gactacytgn gcytcnccyt c
20
      <210> 35
      <211> 30
      <212> DNA
      <213> synthetic
      <400> 35
                                                                           30
      ttaaggtagg actacytcrt aytcytgrct
      <210> 36
      <211> 30
      <212> DNA
      <213> synthetic
      <220><221> misc_feature<222> (27)..(28)<223> n = degeneracy
      <400> 36
45
                                                                           30
      ttaaggtagg actacytcrt aytcytgnga
      <210> 37
50
      <211> 30
      <212> DNA
55
      <213> synthetic
      <220><221> misc_feature<222> (24)..(25)<223> n = degeneracy
       <220><221> misc_feature<222> (27)..(28)<223> n = degeneracy
60
```

	ttaagg	tagg actacrttrw artcngtncc	30
5	<210>	38	
	<211>	16	
	<212>	DNA	
10	<213>	synthetic	
	<400>	38	
15	gatete	tagc actgct	16
	<210>	39	
20	<211>	17	
20	<212>	DNA	
	<213>	synthetic	
25	<400>	39	
25 mm	ccttaa	ggta ggactac	17
30	<210>	40	
	<211>	27	
3 4 4	<212>	DNA	
	<213>	synthetic	
	<400>	40	
40	tatcgc	agcg atccacttct cctctgc	27
	<210>	41	
45	<211>	27	
50	<212>	DNA	
	<213>	synthetic	
	<400>	41	
	cttaaa	.cggc aacctgagaa ggcttgg	27
55	<210>	42	
60	<211>	28	
	<212>	DNA	
	<213>	synthetic	

	<400>	42			
	ctatcto	ctt ctgcttgaaa aca	aggagg		28
5	<210>	43			
	<211>	27			
10	<212>	DNA			
	<213>	synthetic			
15	<400>	43			
13	acaaggg	gaac agctcgtcga tat	cgcg		27
20	<210>	44			
20	<211>	32			
	<212>	DNA			
25	<213>	synthetic			
	<400>	44			
25 May 17 19 19 19 19 19 19 19 19 19 19 19 19 19	taacga	attc ggtgcagaca tag	ggcgaact ac		32
	<210>	45			
<b>3</b> □	<211>	33			
40	<212>	DNA			
	<213>	synthetic			
40	<400>	45			
	cggtgt	cgac tcaggaaaac cad	cctctcaa gcg	3	33
45	<210>	46			
	<211>	37			
50	<212>	DNA			
	<213>	synthetic			
	<400>	46			
55	cacagg	aaac agaccatggg tgo	cagacata gg	cgaac	37
	<210>	47			
60	<211>	1017			
	<212>	DNA			

<213> Ave 47 <400> atgggtgcag acataggcga actactcgag agagaagaag ttgaacttga gtacttctcc 5 120 gggagaaaaa tagctattga tgcttttaac actctttacc agttcatatc tatcataagg caacctgacg gcactccttt gaaggattct cagggtagaa tgacctcaca cctctccggc 180 10 atcctgtacc gcgtgtcaaa catgatcgag gttggaatga gacccatttt cgttttcgat 240 300 360 gcagaggaga agtggatcgc tgcgatagag agaggagaga agtacgcaaa gaagtacgct 15 caggcagcgg cgagggttga tgaatacatc gtcgagtcgt caaagaagct gcttgagtat 420 480 atgggagttc catgggttca ggcgccgagt gagggagagg cacaggctgc atacatggca 20 540 gcgaagggcg atgtagattt tactggctcg caggattacg actcgcttct cttcggcagc ccaaagcttg caagaaatct cgcgattact ggaaagagga agctgcccgg aaagaatgtt 600 tacgttgagg tcaaaccaga gataatagac ttaaacggca acctgagaag gcttggaata 660 25 3 3 3 3 4 720 acaagggaac agctcgtcga tatcgcgttg ctcgtgggaa cggactacaa cgaaggagtg aagggcgttg gggtcaagaa ggcctacaag tacataaaaa cctacggaga tgttttcaaa 780 840 gctctcaagg ccttaaaggt agagcaggag aacatagagg agataagaaa cttcttcctg 900 aacccgcctg ttacgaacaa ctacagcctc cacttcggaa agccagacga tgagaagatt 960 atcgagttcc tgtgtgaaga gcacgacttt agcaaggata gggtagagaa ggccgttgag 1017 aagctgaaag caggaatgca agcctcgcaa tcaacgcttg agaggtggtt ttcctga <210> 48 <211> 337 PRT <212> 45 <213> Ave <400> 48 Met Gly Ala Asp Ile Gly Glu Leu Leu Glu Arg Glu Glu Val Glu Leu 50 Glu Tyr Phe Ser Gly Arg Lys Ile Ala Ile Asp Ala Phe Asn Thr Leu 55 Tyr Gln Phe Ile Ser Ile Ile Arg Gln Pro Asp Gly Thr Pro Leu Lys Asp Ser Gln Gly Arg Met Thr Ser His Leu Ser Gly Ile Leu Tyr Arg

60

Val Ser Asn Met Ile Glu Val Gly Met Arg Pro Ile Phe Val Phe Asp

60

	65					70					75					80
	Gly	Glu	Pro	Pro	Val 85	Phe	Lys	Gln	Lys	Glu 90	Ile	Glu	Glu	Arg	Lys 95	Glu
5	Arg	Arg	Ala	Glu 100	Ala	Glu	Glu	Lys	Trp 105	Ile	Ala	Ala	Ile	Glu 110	Arg	Gly
10	Glu	Lys	Tyr 115	Ala	Lys	Lys	Tyr	Ala 120	Gln	Ala	Ala	Ala	Arg 125	Val	Asp	Glu
	Tyr	Ile 130	Val	Glu	Ser	Ser	Lys 135	Lys	Leu	Leu	Glu	Tyr 140	Met	Gly	Val	Pro
15	Trp 145	Val	Gln	Ala	Pro	Ser 150	Glu	Gly	Glu	Ala	Gln 155	Ala	Ala	Tyr	Met	Ala 160
20	Ala	Lys	Gly	Asp	Val 165	Asp	Phe	Thr	Gly	Ser 170	Gln	Asp	Tyr	Asp	Ser 175	Leu
20	Leu	Phe	Gly	Ser 180	Pro	Lys	Leu	Ala	Arg 185	Asn	Leu	Ala	Ile	Thr 190	Gly	Lys
25	Arg	Lys	Leu 195	Pro	Gly	Lys	Asn	Val 200	Tyr	Val	Glu	Val	Lys 205	Pro	Glu	Ile
The state of the s	Ile	Asp 210	Leu	Asn	Gly	Asn	Leu 215	Arg	Arg	Leu	Gly	Ile 220	Thr	Arg	Glu	Gln
30	Leu 225	Val	Asp	Ile	Ala	Leu 230	Leu	Val	Gly	Thr	Asp 235	Tyr	Asn	Glu	Gly	Val 240
35	Lys	Gly	Val	Gly	Val 245	Lys	Lys	Ala	Tyr	Lys 250	Tyr	Ile	Lys	Thr	Tyr 255	Gly
	Asp	Val	Phe	Lys 260	Ala	Leu	Lys	Ala	Leu 265	Lys	Val	Glu	Gln	Glu 270	Asn	Ile
40	Glu	Glu	Ile 275	Arg	Asn	Phe	Phe	Leu 280	Asn	Pro	Pro	Val	Thr 285	Asn	Asn	Tyr
	Ser	Leu 290	His	Phe	Gly	Lys	Pro 295	Asp	Asp	Glu	Lys	Ile 300	Ile	Glu	Phe	Leu
45	Cys 305	Glu	Glu	His	Asp	Phe 310	Ser	Lys	Asp	Arg	Val 315	Glu	Lys	Ala	Val	Glu 320
50	Lys	Leu	Lys	Ala	Gly 325	Met	Gln	Ala	Ser	Gln 330	Ser	Thr	Leu	Glu	Arg 335	Trp
50	Phe	Ser														
55	<21		49													
	<21 <21		53 DNA													
60	<21	.3>	synt	heti	С											
		_														

<400> 49

```
cccgtctcgc tggtgaaaag aaaaaccacc ctggcgccca atacgcaaac cgc
                                                                          53
      <210> 50
5
      <211> 26
      <212> DNA
      <213> synthetic
10
      <220><221> misc_feature<222> (22)..(23)<223> misc. feature
      <220><221> misc_feature<222> (1)..(2)<223> misc. feature
15
      <220><221> misc_feature<222> (2)..(3)<223> misc. feature
      <400> 50
                                                                           26
      tntnccagag cctaatttgc cagtna
20
      <210> 51
      <211> 26
      <212> DNA
      <213> synthetic
      <220><221> misc_feature<222> (1)..(2)<223> misc. feature
      <220><221> misc_feature<222> (2)..(3)<223> Misc. Feature
      <220><221> misc_feature<222> (22)..(23)<223> Misc. Feature
      <400> 51
                                                                           26
      tntnccagag cctaatttgc cagtna
      <210> 52
      <211> 24
45
      <212> DNA
      <213> synthetic
      <220><221> misc_feature<222> (22)..(23)<223> misc. feature
50
       <400> 52
                                                                           24
       ttccagagcc taatttgcca gtna
55
       <210> 53
       <211>
             24
60
```

<212> DNA

```
<213> synthetic
      <220><221> misc_feature<222> (22)..(23)<223> misc. feature
      <400> 53
5
                                                                          24
      ttccagagcc taatttgcca gtna
      <210> 54
10
      <211> 25
      <212> DNA
15
      <213> synthetic
      <400> 54
                                                                          25
      cttaccaacg ctaacgagcg tcttg
20
      <210> 55
      <211> 14
      <212> DNA
      <213> synthetic
      <400> 55
                                                                          14
      gctcccgcag acac
      <210> 56
      <211> 15
      <212> DNA
      <213> synthetic
      <220><221> misc_feature<222> (1)..(2)<223> misc. feature
45
      <400> 56
                                                                           15
      tntacgccac cagct
50
       <210> 57
       <211> 12
      <212> DNA
55
       <213> synthetic
       <400> 57
60
                                                                           12
       cgctgtctcg ct
```

```
<210> 58
       <211> 19
 5
      <212>
             DNA
       <213> synthetic
       <400> 58
10
                                                                              19
       gctcaaggca ctcttgccc
       <210> 59
15
       <211>
             63
       <212>
             DNA
       <213> synthetic
20
       <400> 59
       atgactgaat ataaacttgt ggtagttgga gctggtggcg taggcaagag tgccttgacg
                                                                              60
30 30 3 3 40
                                                                              63
       ata
       <210> 60
       <211>
             45
       <212> DNA
       <213> synthetic
       <400> 60
                                                                              45
       ttttttttta attaggctct ggaaagacgc tcgtgaaacg agcgt
       <210> 61
       <211>
              14
45
       <212> DNA
       <213>
             synthetic
50
       <400> 61
                                                                              14
       cttcggagtt tggg
       <210> 62
55
       <211>
              16
              DNA
       <212>
60
```

<213> synthetic

```
<220><221> misc_feature<222> (1)..(2)<223> misc. feature
      <400> 62
                                                                           16
5
      ancttcggag tttggg
      <210> 63
      <211> 16
10
      <212> DNA
      <213> synthetic
15
      <220><221> misc_feature<222> (1)..(2)<223> misc. feature
      <400> 63
                                                                           16
      cncttcggag tttggg
20
      <210> 64
      <211> 16
3 3 3 4
      <212> DNA
      <213> synthetic
       <220><221> misc_feature<222> (1)..(2)<223> misc. feature
       <400> 64
                                                                           16
      gncttcggag tttggg
       <210> 65
       <211> 16
       <212> DNA
       <213> synthetic
       <220><221> misc_feature<222> (1)..(2)<223> misc. feature
45
       <400> 65
                                                                            16
       tncttcggag tttggg
50
       <210> 66
       <211> 25
55
       <212> DNA
       <213> synthetic
60
       <400> 66
                                                                            25
       gggttgtgga gtgagtgttc aagta
```

	<210>	67								
5	<211>	27								
	<212>	DNA								
	<213>	synt:	hetic							
10	<400>	67								
	ccatcc	taat	acgactcact	atagggc				27		
15	<210>	68								
	<211>	21								
20	<212> DNA									
20	<213> synthetic									
	<400>	68								
25 30 30 White with 1978 375 375 375 375 375 375 375 375 375 375	ctcata	cagt	tacttgtctt	С				21		
	<210>	69								
30	<211>	489								
	<212>	RNA								
	<213>	Homo	sapiens							
	<400>	69								
and the state of t	gaacuc	acua	uagggcucga	gcggccgccc	gggcaggucc	gccaccaaaa	ugcagauuuu	60		
40	cgugaa	aacc	cuuacgggga	agaccaucac	ccucgagguu	gaacccucgg	auacgauaga	120		
	aaaugu	aaag	gccaagaucc	aggauaagga	aggaauuccu	ccugaucagc	agagacugau	180		
[så	cuuugc	uggc	aagcagcugg	aagauggacg	uacuuugucu	gacuacaaua	uucaaaagga	240		
45	gucuac	ucuu	caucuugugu	ugagacuucg	ugguggugcu	aagaaaagga	agaagaaguc	300		
	uuacac	cacu	cccaagaaga	auaagcacaa	gagaaagaag	guuaagcugg	cuguccugaa	360		
50	auauua	uaag	guggaugaga	auggcaaaau	uagucgccuu	cgucgagagu	gcccuucuga	420		
*	ugaaug	juggu	gcuggggugu	uuauggcaag	ucacuuugac	agacauuauu	guggcaaaug	480		
	uugucu	ıgac						489		
55	0.1.0	<b>5</b> 0								
	<210>	70								
60	<211>	52								
	<212>	DNA								

	<213>	synt	hetic					
	<400>	70						
5	ggaatao	gac	tcactatagg	gaaagtctct	gccgcccttc	tgtgcctgct	gc	52
	<210>	71						
10	<211>	52						
	<212>	DNA						
	<213>	synt	thetic					
15	<400>	71						
	ggaatao	cgac	tcactatagg	gaaagtctct	gccgcccttc	tgtgcctgct	gc	52
20	<210>	72						
	<211>	647						
25		RNA						
	<213>	synt	chetic					
2 5 The state of t	<400>	72						
30	gggaaag	gucu	cugeegeecu	ucugugccug	cugcucauag	cagccaccuu	cauuccccaa	60
	gggcuc	gcuc	agccagaugc	aaucaaugcc	ccagucaccu	gcuguuauaa	cuucaccaau	120
35	aggaag	aucu	cagugcagag	gcucgcgagc	uauagaagaa	ucaccagcag	caaguguccc	180
	aaagaa	gcug	ugaucuucaa	gaccauugug	gccaaggaga	ucugugcuga	ccccaagcag	240
4 0	aagugg	guuc	aggauuccau	ggaccaccug	gacaagcaaa	cccaaacucc	gaagacuuga	300
40%	acacuc	acuc	cacaacccaa	gaaucugcag	cuaacuuauu	uuccccuagc	uuuccccaga	360
<b>1</b>	cacccu	guuu	uauuuuauua	uaaugaauuu	uguuuguuga	ugugaaacau	uaugccuuaa	420
45	guaaug	uuaa	uucuuauuua	aguuauugau	guuuuaaguu	uaucuuucau	gguacuagug	480
	uuuuuu	agau	acagagacuu	ggggaaauug	cuuuuccucu	ugaaccacag	uucuaccccu	540
50	gggaug	uuuu	gagggucuuu	gcaagaauca	uuaauacaaa	gaauuuuuuu	uaacauucca	600
30	augcau	ugcu	aaaauauuau	uguggaaaug	aauauuuugu	aacuauu		647
	<210>	73						
55	<211>	16						
	<212>	DNA						
60	<213>	syn	thetic					

<400> 73

	ttcttcggag tttggg	16
5	<210> 74	
	<211> 26	
	<212> DNA	
10	<213> synthetic	
	<400> 74	
15	ccgtcacgcc tccttcggag tttggg	26
	<210> 75	
0.0	<211> 24	
20	<212> DNA	
	<213> synthetic	
25	<400> 75	
25 30 3 35 35 35 35 35 35 35 35 35 35 35 35 3	aacccaaact ccgaaggagg cgtg	24
30 4	<210> 76	
104 14 14	<211> 29	
2 년	<212> DNA	
8:	<213> synthetic	
Anna Anna Anna Anna Anna Anna Anna Anna	<400> 76	
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	gcgcagtgag aatgaggagg cgtgacggt	29
100 Marie 100 Ma	<210> 77	
45	<211> 17	
	<212> DNA	
50	<213> synthetic	
	<220><221> misc_feature<222> (1)(2)<223> misc. feature	
	<400> 77	
55	cnctcattct cagtgcg	17
	<210> 78	
60	<211> 30	
	<212> DNA	

```
<213> synthetic
                                  <400> 78
                                                                                                                                                                                                                                                                                                                                                                                                     30
    5
                                 aacgaggcgc acctttacat tttctatcgt
                                 <210> 79
                                 <211> 24
10
                                  <212> DNA
                                  <213> synthetic
15
                                  <400> 79
                                                                                                                                                                                                                                                                                                                                                                                                      24
                                  ccttccttat cctggatctt ggca
20
                                  <210> 80
                                  <211> 24
25
                                  <212> DNA
                                  <213> synthetic
                                   <400> 80
30 35
                                                                                                                                                                                                                                                                                                                                                                                                       24
                                  acgatagaaa atgtaaaggt gcgc
                                  <210> 81
     The state of the s
                                  <211> 29
                                   <212> DNA
 40
                                   <213> synthetic
       <400> 81
                                                                                                                                                                                                                                                                                                                                                                                                        29
        i.e.l
                                    cggaagaagc aagtggtgcg cctcgttaa
 45
                                    <210> 82
                                    <211> 16
 50
                                    <212> DNA
                                    <213> synthetic
                                    <220><221> misc_feature<222> (1)..(2)<223> misc. feature
  55
                                    <400> 82
                                                                                                                                                                                                                                                                                                                                                                                                         16
                                     gncacttgct tcctcc
   60
```

<210> 83

	<211>	14								
	<212>	DNA								
5	<213>	synthetic								
	<400>	83								
10	gctcccgcag acac 14									
	<210>	84								
15	<211>	30								
12	<212>	DNA								
	<213>	synthetic								
20	<400>	84								
	caaaga	aaag ctgcgtgatg atgaaatcgc	30							
25	<210>	85								
30	<211>	50								
30	<212>	DNA								
30 ·	<213>	synthetic								
- Marian	<400>	85								
35	gaaggt	gtct gegggageeg attteateat caegeagett ttetttgagg	50							
And South And South And										